Linear Algebra

M.Tech. C.S.

Syllabus

Matrices and System of Equations: System of Linear Equations, Row Echelon Form, Matrix Arithmetic, Matrix Algebra, Elementary Matrices, Partitioned Matrices, Determinant and its properties.

Vector Spaces: Definition and Examples, Subspaces, Linear Independence, Basis and Dimension, Change of Basis, Row Space, Column Space, Null space

Inner product spaces: The Euclidean dot product, Orthogonal Subspaces, Least Squares Problems, Orthonormal Sets, The Gram-Schmidt Orthogonalization Process, Orthogonal Polynomials

Linear Transformations: Definition and Examples, Matrix Representations of Linear Transformations, Similarity

Eigenvalues and Eigenvectors: System of Linear Differential Equations, Diagonalization, Hermitian Matrices, Singular Value Decomposition.

Quadratic Forms: Classifiction and characterisations, Optimisation of quadratic forms.

Algorithms: Gaussian Elimination with different Pivoting Strategies; Matrix Norms and Condition Numbers; Orthogonal Transformations; The Eigenvalue Problem; Least Squares Problems.

References:

- *Linear Algebra Done Right* by Sheldon Axler, Springer
- *Linear Algebra with Applications* by Steve Leon, Pearson Global edition.
- *Introduction to Linear Algebra* by Gilbert Strang

1.1 Systems of Linear Equations

A *linear equation in n unknowns* is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where $a_1, a_2, \ldots, a_n$ and b are real numbers and $x_1, x_2, \ldots, x_n$ are variables. A *linear system* of m equations in n unknowns is then a system of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1}$$

where the a_{ij}'s and the b_i's are all real numbers. We will refer to systems of the form (1) as $m \times n$ linear systems. The following are examples of linear systems:

(a) $$\begin{aligned} x_1 + 2x_2 &= 5 \\ 2x_1 + 3x_2 &= 8 \end{aligned}$$

(b) $$\begin{aligned} x_1 - x_2 + x_3 &= 2 \\ 2x_1 + x_2 - x_3 &= 4 \end{aligned}$$

(c) $$\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 1 \\ x_1 &= 4 \end{aligned}$$

Definition Two systems of equations involving the same variables are said to be **equivalent** if they have the same solution set.

Clearly, if we interchange the order in which two equations of a system are written, this will have no effect on the solution set. The reordered system will be equivalent to the original system. For example, the systems

$$\begin{aligned} x_1 + 2x_2 &= 4 \\ 3x_1 - x_2 &= 2 \\ 4x_1 + x_2 &= 6 \end{aligned} \qquad \text{and} \qquad \begin{aligned} 4x_1 + x_2 &= 6 \\ 3x_1 - x_2 &= 2 \\ x_1 + 2x_2 &= 4 \end{aligned}$$

both involve the same three equations and, consequently, they must have the same solution set.

If one equation of a system is multiplied through by a nonzero real number, this will have no effect on the solution set, and the new system will be equivalent to the original system. For example, the systems

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ -2x_1 - x_2 + 4x_3 &= 1 \end{aligned} \qquad \text{and} \qquad \begin{aligned} 2x_1 + 2x_2 + 2x_3 &= 6 \\ -2x_1 - x_2 + 4x_3 &= 1 \end{aligned}$$

are equivalent.

To summarize, there are three operations that can be used on a system to obtain an equivalent system:

I. The order in which any two equations are written may be interchanged.

II. Both sides of an equation may be multiplied by the same nonzero real number.

III. A multiple of one equation may be added to (or subtracted from) another.

Given a system of equations, we may use these operations to obtain an equivalent system that is easier to solve.

$n \times n$ Systems

Let us restrict ourselves to $n \times n$ systems for the remainder of this section. We will show that if an $n \times n$ system has exactly one solution, then operations **I** and **III** can be used to obtain an equivalent "strictly triangular system."

Definition

A system is said to be in **strict triangular form** if, in the kth equation, the coefficients of the first $k-1$ variables are all zero and the coefficient of x_k is nonzero $(k = 1, \ldots, n)$.

EXAMPLE 1 The system

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

If we attach to the coefficient matrix an additional column whose entries are the numbers on the right-hand side of the system, we obtain the new matrix

$$\left(\begin{array}{rrr|r} 1 & 2 & 1 & 3 \\ 3 & -1 & -3 & -1 \\ 2 & 3 & 1 & 4 \end{array}\right)$$

We will refer to this new matrix as the *augmented matrix*. In general, when an $m \times r$ matrix B is attached to an $m \times n$ matrix A in this way, the augmented matrix is denoted by $(A|B)$. Thus, if

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2r} \\ \vdots & & & \\ b_{m1} & b_{m2} & \cdots & b_{mr} \end{pmatrix}$$

then

$$(A|B) = \left(\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1n} & b_{11} & \cdots & b_{1r} \\ \vdots & & & \vdots & & \\ a_{m1} & \cdots & a_{mn} & b_{m1} & \cdots & b_{mr} \end{array}\right)$$

With each system of equations, we may associate an augmented matrix of the form

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right)$$

ELEMENTARY ROW OPERATIONS

1. (Replacement) Replace one row by the sum of itself and a multiple of another row.[1]
2. (Interchange) Interchange two rows.
3. (Scaling) Multiply all entries in a row by a nonzero constant.

Row operations can be applied to any matrix, not merely to one that arises as the augmented matrix of a linear system. Two matrices are called **row equivalent** if there is a sequence of elementary row operations that transforms one matrix into the other.

Example

Solve the system of equations

$$\begin{aligned} x + 2y + 3z &= 6 \\ 2x - 3y + 2z &= 14 \\ 3x + y - z &= -2 \end{aligned}$$

Matrix notation:

$$\begin{aligned} x + 2y + 3z &= 6 \\ 2x - 3y + 2z &= 14 \\ 3x + y - z &= -2 \end{aligned} \quad \overset{\text{becomes}}{\rightsquigarrow} \quad \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right)$$

This is an **augmented matrix**.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right)$$

Goal: we want our elimination method to eventually produce a system of equations like

$$\begin{aligned} x \qquad\qquad &= A \\ y \qquad &= B \\ z &= C \end{aligned} \qquad \text{or in matrix form,} \qquad \left(\begin{array}{ccc|c} 1 & 0 & 0 & A \\ 0 & 1 & 0 & B \\ 0 & 0 & 1 & C \end{array}\right)$$

So we need to do row operations that make the start matrix look like the end one.

Row Operations

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ \boxed{2} & -3 & 2 & 14 \\ \boxed{3} & 1 & -1 & -2 \end{array}\right) \xrightarrow{R_2 = R_2 - 2R_1} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 3 & 1 & -1 & -2 \end{array}\right)$$

We want these to be zero.
So we subract multiples of the first row.

$$\xrightarrow{R_3 = R_3 - 3R_1} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 0 & -5 & -10 & -20 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & \boxed{2} & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 0 & \boxed{-5} & -10 & -20 \end{array}\right) \xrightarrow{R_2 \longleftrightarrow R_3} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -5 & -10 & -20 \\ 0 & -7 & -4 & 2 \end{array}\right)$$

We want these to be zero.

It would be nice if this were a 1.
We could divide by -7, but that
would produce ugly fractions.

Let's swap the last two rows first.

$$\xrightarrow{R_2 = R_2 \div -5} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & -7 & -4 & 2 \end{array}\right)$$

$$\xrightarrow{R_1 = R_1 - 2R_2} \left(\begin{array}{ccc|c} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 4 \\ 0 & -7 & -4 & 2 \end{array}\right)$$

$$\xrightarrow{R_3 = R_3 + 7R_2} \left(\begin{array}{ccc|c} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 10 & 30 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 0 & \boxed{\begin{matrix}-1\\ 2\end{matrix}} & -2 \\ 0 & 1 & & 4 \\ 0 & 0 & 10 & 30 \end{array}\right) \xrightarrow{R_3 = R_3 \div 10} \left(\begin{array}{ccc|c} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

We want these to be zero.

Let's make this a 1 first.

$$\xrightarrow{R_1 = R_1 + R_3} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

$$\xrightarrow{R_2 = R_2 - 2R_3} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

$$\xrightarrow{\text{translates into}} \begin{aligned} x \quad &= 1 \\ y \quad &= -2 \\ z &= 3 \end{aligned}$$

Check:

$$\begin{aligned} x + 2y + 3z &= 6 \\ 2x - 3y + 2z &= 14 \\ 3x + y - z &= -2 \end{aligned} \xrightarrow{\text{substitute solution}} \begin{aligned} 1 + 2\cdot(-2) + 3\cdot 3 &= 6 \\ 2\cdot 1 - 3\cdot(-2) + 2\cdot 3 &= 14 \\ 3\cdot 1 + (-2) - 3 &= -2 \end{aligned}$$

1.2 Row Echelon Form

A rectangular matrix is in **echelon form** (or **row echelon form**) if it has the following three properties:

1. All nonzero rows are above any rows of all zeros.
2. Each leading entry of a row is in a column to the right of the leading entry of the row above it.
3. All entries in a column below a leading entry are zeros.

If a matrix in echelon form satisfies the following additional conditions, then it is in **reduced echelon form** (or **reduced row echelon form**):

4. The leading entry in each nonzero row is 1.
5. Each leading 1 is the only nonzero entry in its column.

EXAMPLE 1 The following matrices are in echelon form. The leading entries (■) may have any nonzero value; the starred entries (∗) may have any value (including zero).

$$\begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * \end{bmatrix}$$

The following matrices are in reduced echelon form because the leading entries are 1's, and there are 0's below *and above* each leading 1.

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

■

The "triangular" matrices

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

are in echelon form. In fact, the second matrix is in reduced echelon form.

THEOREM 1

Uniqueness of the Reduced Echelon Form

Each matrix is row equivalent to one and only one reduced echelon matrix.

Proof.: Home task.

DEFINITION

A **pivot position** in a matrix A is a location in A that corresponds to a leading 1 in the reduced echelon form of A. A **pivot column** is a column of A that contains a pivot position.

EXAMPLE 2 Row reduce the matrix A below to echelon form, and locate the pivot columns of A.

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

SOLUTION Use the same basic strategy as in Section 1.1. The top of the leftmost nonzero column is the first pivot position. A nonzero entry, or *pivot*, must be placed in this position. A good choice is to interchange rows 1 and 4 (because the mental computations in the next step will not involve fractions).

$$\begin{array}{c}
\text{Pivot} \\
\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \\
\text{Pivot column}
\end{array}$$

Create zeros below the pivot, 1, by adding multiples of the first row to the rows below, and obtain matrix (1) below. The pivot position in the second row must be as far left as possible—namely, in the second column. Choose the 2 in this position as the next pivot.

$$\begin{array}{c}
\text{Pivot} \\
\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \\
\text{Next pivot column}
\end{array} \tag{1}$$

Add $-5/2$ times row 2 to row 3, and add $3/2$ times row 2 to row 4.

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix} \tag{2}$$

The matrix in (2) is different from any encountered in Section 1.1. There is no way to create a leading entry in column 3! (We can't use row 1 or 2 because doing so would destroy the echelon arrangement of the leading entries already produced.) However, if we interchange rows 3 and 4, we can produce a leading entry in column 4.

Pivot

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{General form:} \quad \begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Pivot columns

The matrix is in echelon form and thus reveals that columns 1, 2, and 4 of A are pivot columns.

Pivot positions

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

Pivot columns

Equivalent Systems

Given an $m \times n$ linear system $A\mathbf{x} = \mathbf{b}$, we can obtain an equivalent system by multiplying both sides of the equation by a nonsingular $m \times m$ matrix M:

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

$$MA\mathbf{x} = M\mathbf{b} \tag{2}$$

Clearly, any solution of (1) will also be a solution of (2). On the other hand, if $\hat{\mathbf{x}}$ is a solution of (2), then

$$M^{-1}(MA\hat{\mathbf{x}}) = M^{-1}(M\mathbf{b})$$

$$A\hat{\mathbf{x}} = \mathbf{b}$$

and it follows that the two systems are equivalent.

To transform the system $A\mathbf{x} = \mathbf{b}$ to a simpler form that is easier to solve, we can apply a sequence of nonsingular matrices $E_1, \ldots, E_k$ to both sides of the equation. The new system will then be of the form

$$U\mathbf{x} = \mathbf{c}$$

where $U = E_k \cdots E_1 A$ and $\mathbf{c} = E_k \cdots E_2 E_1 \mathbf{b}$. The transformed system will be equivalent to the original, provided that $M = E_k \cdots E_1$ is nonsingular. However, M is nonsingular, since it is a product of nonsingular matrices.

Elementary Matrices

If we start with the identity matrix I and then perform exactly one elementary row operation, the resulting matrix is called an *elementary* matrix.

There are three types of elementary matrices corresponding to the three types of elementary row operations.

Type I An elementary matrix of type I is a matrix obtained by interchanging two rows of I.

EXAMPLE 1 The matrix
$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
is an elementary matrix of type I since it was obtained by interchanging the first two rows of I. If A is a 3×3 matrix, then

$$E_1A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$AE_1 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{12} & a_{11} & a_{13} \\ a_{22} & a_{21} & a_{23} \\ a_{32} & a_{31} & a_{33} \end{bmatrix}$$

Multiplying A on the left by E_1 interchanges the first and second rows of A. Right multiplication of A by E_1 is equivalent to the elementary column operation of interchanging the first and second columns. ■

Type II An elementary matrix of type II is a matrix obtained by multiplying a row of I by a nonzero constant.

EXAMPLE 2

$$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

is an elementary matrix of type II. If A is a 3×3 matrix, then

$$E_2A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 3a_{31} & 3a_{32} & 3a_{33} \end{bmatrix}$$

$$AE_2 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & 3a_{13} \\ a_{21} & a_{22} & 3a_{23} \\ a_{31} & a_{32} & 3a_{33} \end{bmatrix}$$

Multiplication on the left by E_2 performs the elementary row operation of multiplying the third row by 3, while multiplication on the right by E_2 performs the elementary column operation of multiplying the third column by 3. ■

Type III An elementary matrix of type III is a matrix obtained from I by adding a multiple of one row to another row.

EXAMPLE 3

$$E_3 = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is an elementary matrix of type III. If A is a 3×3 matrix, then

$$E_3 A = \begin{bmatrix} a_{11} + 3a_{31} & a_{12} + 3a_{32} & a_{13} + 3a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$A E_3 = \begin{bmatrix} a_{11} & a_{12} & 3a_{11} + a_{13} \\ a_{21} & a_{22} & 3a_{21} + a_{23} \\ a_{31} & a_{32} & 3a_{31} + a_{33} \end{bmatrix}$$

Multiplication on the left by E_3 adds 3 times the third row to the first row. Multiplication on the right adds 3 times the first column to the third column. ■

Theorem 1.5.1 If E is an elementary matrix, then E is nonsingular and E^{-1} is an elementary matrix of the same type.

Proof If E is the elementary matrix of type I formed from I by interchanging the ith and jth rows, then E can be transformed back into I by interchanging these same rows again. Therefore, $EE = I$ and hence E is its own inverse. If E is the elementary matrix of type II formed by multiplying the ith row of I by a nonzero scalar α, then E can be transformed into the identity matrix by multiplying either its ith row or its ith column by $1/\alpha$. Thus,

$$E^{-1} = \begin{array}{cl} \left[\begin{array}{ccccccc} 1 & & & & & & \\ & \ddots & & & & O & \\ & & 1 & & & & \\ & & & 1/\alpha & & & \\ & & & & 1 & & \\ & O & & & & \ddots & \\ & & & & & & 1 \end{array}\right] & \begin{array}{l} \\ \\ \\ i\text{th row} \\ \\ \\ \\ \end{array} \end{array}$$

Finally, if E is the elementary matrix of type III formed from I by adding m times the

ith row to the jth row, that is,

$$
E = \begin{bmatrix}
1 & & & & & & \\
\vdots & \ddots & & & & O & \\
0 & \cdots & 1 & & & & \\
\vdots & & & \ddots & & & \\
0 & \cdots & m & \cdots & 1 & & \\
\vdots & & & & & \ddots & \\
0 & \cdots & 0 & \cdots & 0 & \cdots & 1
\end{bmatrix}
\begin{array}{l}
\\ \\ i\text{th row} \\ \\ j\text{th row} \\ \\ \\
\end{array}
$$

then E can be transformed back into I either by subtracting m times the ith row from the jth row or by subtracting m times the jth column from the ith column. Thus,

$$
E^{-1} = \begin{bmatrix}
1 & & & & & & \\
\vdots & \ddots & & & & O & \\
0 & \cdots & 1 & & & & \\
\vdots & & & \ddots & & & \\
0 & \cdots & -m & \cdots & 1 & & \\
\vdots & & & & & \ddots & \\
0 & \cdots & 0 & \cdots & 0 & \cdots & 1
\end{bmatrix}
$$

■

Definition

A matrix B is **row equivalent** to a matrix A if there exists a finite sequence $E_1, E_2, \ldots, E_k$ of elementary matrices such that

$$B = E_k E_{k-1} \cdots E_1 A$$

In other words, B is row equivalent to A if B can be obtained from A by a finite number of row operations. In particular, if two augmented matrices $(A \mid \mathbf{b})$ and $(B \mid \mathbf{c})$ are row equivalent, then $A\mathbf{x} = \mathbf{b}$ and $B\mathbf{x} = \mathbf{c}$ are equivalent systems.

The following properties of row equivalent matrices are easily established:

I. If A is row equivalent to B, then B is row equivalent to A.

II. If A is row equivalent to B, and B is row equivalent to C, then A is row equivalent to C.

Property (I) can be proved using Theorem 1.5.1. The details of the proofs of (I) and (II) are left as an exercise for the reader.

Theorem 1.5.2 Equivalent Conditions for Nonsingularity

Let A be an $n \times n$ matrix. The following are equivalent:

(a) A is nonsingular.

(b) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{0}$.

(c) A is row equivalent to I.

Proof We prove first that statement (a) implies statement (b). If A is nonsingular and $\hat{\mathbf{x}}$ is a solution of $A\mathbf{x} = \mathbf{0}$, then

$$\hat{\mathbf{x}} = I\hat{\mathbf{x}} = (A^{-1}A)\hat{\mathbf{x}} = A^{-1}(A\hat{\mathbf{x}}) = A^{-1}\mathbf{0} = \mathbf{0}$$

Thus, $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Next, we show that statement (b) implies statement (c). If we use elementary row operations, the system can be transformed into the form $U\mathbf{x} = \mathbf{0}$, where U is in row echelon form. If one of the diagonal elements of U were 0, the last row of U would consist entirely of 0's. But then $A\mathbf{x} = \mathbf{0}$ would be equivalent to a system with more unknowns than equations and, hence, by Theorem 1.2.1, would have a nontrivial solution. Thus, U must be a strictly triangular matrix with diagonal elements all equal to 1. It then follows that I is the reduced row echelon form of A and hence A is row equivalent to I.

Finally, we will show that statement (c) implies statement (a). If A is row equivalent to I, there exist elementary matrices $E_1, E_2, \ldots, E_k$ such that

$$A = E_k E_{k-1} \cdots E_1 I = E_k E_{k-1} \cdots E_1$$

But since E_i is invertible, $i = 1, \ldots, k$, the product $E_k E_{k-1} \cdots E_1$ is also invertible. Hence, A is nonsingular and

$$A^{-1} = (E_k E_{k-1} \cdots E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1} \qquad \blacksquare$$

Corollary 1.5.3 The system $A\mathbf{x} = \mathbf{b}$ of n linear equations in n unknowns has a unique solution if and only if A is nonsingular.

Proof If A is nonsingular and $\hat{\mathbf{x}}$ is any solution of $A\mathbf{x} = \mathbf{b}$, then

$$A\hat{\mathbf{x}} = \mathbf{b}$$

Multiplying both sides of this equation by A^{-1}, we see that $\hat{\mathbf{x}}$ must be equal to $A^{-1}\mathbf{b}$.

Conversely, if $A\mathbf{x} = \mathbf{b}$ has a unique solution $\hat{\mathbf{x}}$, then we claim that A cannot be singular. Indeed, if A were singular, then the equation $A\mathbf{x} = \mathbf{0}$ would have a solution $\mathbf{z} \neq \mathbf{0}$. But this would imply that $\mathbf{y} = \hat{\mathbf{x}} + \mathbf{z}$ is a second solution of $A\mathbf{x} = \mathbf{b}$, since

$$A\mathbf{y} = A(\hat{\mathbf{x}} + \mathbf{z}) = A\hat{\mathbf{x}} + A\mathbf{z} = \mathbf{b} + \mathbf{0} = \mathbf{b}$$

Therefore, if $A\mathbf{x} = \mathbf{b}$ has a unique solution, then A must be nonsingular. ■

If A is nonsingular, then A is row equivalent to I and hence there exist elementary matrices $E_1, \ldots, E_k$ such that

$$E_k E_{k-1} \cdots E_1 A = I$$

Multiplying both sides of this equation on the right by A^{-1}, we obtain

$$E_k E_{k-1} \cdots E_1 I = A^{-1}$$

Thus, the same series of elementary row operations that transforms a nonsingular matrix A into I will transform I into A^{-1}. This gives us a method for computing A^{-1}. If we augment A by I and perform the elementary row operations that transform A into I on the augmented matrix, then I will be transformed into A^{-1}. That is, the reduced row echelon form of the augmented matrix $(A \mid I)$ will be $\left(I \mid A^{-1}\right)$.

EXAMPLE 4 Compute A^{-1} if

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

Solution

$$\left[\begin{array}{rrr|rrr} 1 & 4 & 3 & 1 & 0 & 0 \\ -1 & -2 & 0 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rrr|rrr} 1 & 4 & 3 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & 1 & 0 \\ 0 & -6 & -3 & -2 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{rrr|rrr} 1 & 4 & 3 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & 1 & 0 \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rrr|rrr} 1 & 4 & 0 & \frac{1}{2} & -\frac{3}{2} & -\frac{1}{2} \\ 0 & 2 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 2 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{array}\right]$$

Thus,

$$A^{-1} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{bmatrix}$$ ■

EXAMPLE 5 Solve the system

$$\begin{aligned} x_1 + 4x_2 + 3x_3 &= 12 \\ -x_1 - 2x_2 \phantom{{}+3x_3} &= -12 \\ 2x_1 + 2x_2 + 3x_3 &= 8 \end{aligned}$$

Solution

The coefficient matrix of this system is the matrix A of the last example. The solution of the system is then

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 12 \\ -12 \\ 8 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ -\frac{8}{3} \end{bmatrix}$$

■

Triangular Factorization

If an $n \times n$ matrix A can be reduced to strict upper triangular form using only row operation III, then it is possible to represent the reduction process in terms of a matrix factorization. We illustrate how this is done in the next example.

EXAMPLE 6 Let

$$A = \begin{bmatrix} 2 & 4 & 2 \\ 1 & 5 & 2 \\ 4 & -1 & 9 \end{bmatrix}$$

and let us use only row operation III to carry out the reduction process. At the first step, we subtract $\frac{1}{2}$ times the first row from the second and then we subtract twice the first row from the third.

$$\begin{bmatrix} 2 & 4 & 2 \\ 1 & 5 & 2 \\ 4 & -1 & 9 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & -9 & 5 \end{bmatrix}$$

To keep track of the multiples of the first row that were subtracted, we set $l_{21} = \frac{1}{2}$ and $l_{31} = 2$. We complete the elimination process by eliminating the -9 in the $(3, 2)$ position:

$$\begin{bmatrix} 2 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & -9 & 5 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 8 \end{bmatrix}$$

Let $l_{32} = -3$, the multiple of the second row subtracted from the third row. If we call the resulting matrix U and set

$$L = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 2 & -3 & 1 \end{bmatrix}$$

then it is easily verified that

$$LU = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 2 & -3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 2 \\ 1 & 5 & 2 \\ 4 & -1 & 9 \end{bmatrix} = A$$

The matrix L in the previous example is lower triangular with 1's on the diagonal. We say that L is *unit lower triangular*. The factorization of the matrix A into a product of a unit lower triangular matrix L times a strictly upper triangular matrix U is often referred to as an *LU factorization*.

$$E_3E_2E_1A = U \tag{3}$$

where

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix}$$

correspond to the row operations in the reduction process. Since each of the elementary matrices is nonsingular, we can multiply equation (3) by their inverses.

$$A = E_1^{-1}E_2^{-1}E_3^{-1}U$$

[We multiply in reverse order because $(E_3E_2E_1)^{-1} = E_1^{-1}E_2^{-1}E_3^{-1}$.] However, when the inverses are multiplied in this order, the multipliers l_{21}, l_{31}, l_{32} fill in below the diagonal in the product:

$$E_1^{-1}E_2^{-1}E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix} = L$$

In general, if an $n \times n$ matrix A can be reduced to strict upper triangular form using only row operation III, then A has an LU factorization. The matrix L is unit lower triangular, and if $i > j$, then l_{ij} is the multiple of the jth row subtracted from the ith row during the reduction process.

1.6 Partitioned Matrices

Often it is useful to think of a matrix as being composed of a number of submatrices. A matrix C can be partitioned into smaller matrices by drawing horizontal lines between the rows and vertical lines between the columns. The smaller matrices are often referred to as *blocks*. For example, let

$$C = \begin{pmatrix} 1 & -2 & 4 & 1 & 3 \\ 2 & 1 & 1 & 1 & 1 \\ 3 & 3 & 2 & -1 & 2 \\ 4 & 6 & 2 & 2 & 4 \end{pmatrix}$$

If lines are drawn between the second and third rows and between the third and fourth columns, then C will be divided into four submatrices, C_{11}, C_{12}, C_{21}, and C_{22}:

$$\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = \left(\begin{array}{ccc|cc} 1 & -2 & 4 & 1 & 3 \\ 2 & 1 & 1 & 1 & 1 \\ \hline 3 & 3 & 2 & -1 & 2 \\ 4 & 6 & 2 & 2 & 4 \end{array}\right)$$

One useful way of partitioning a matrix is into columns. For example, if

$$B = \begin{pmatrix} -1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 4 & 1 \end{pmatrix}$$

then we can partition B into three column submatrices:

$$B = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = \left(\begin{array}{r|r|r} -1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 4 & 1 \end{array}\right)$$

Suppose that we are given a matrix A with three columns; then the product AB can be viewed as a block multiplication. Each block of B is multiplied by A, and the result is a matrix with three blocks: $A\mathbf{b}_1$, $A\mathbf{b}_2$, and $A\mathbf{b}_3$; that is,

$$AB = A(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = (A\mathbf{b}_1, A\mathbf{b}_2, A\mathbf{b}_3)$$

For example, if

$$A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 1 & -2 \end{pmatrix}$$

then

$$A\mathbf{b}_1 = \begin{bmatrix} 6 \\ -2 \end{bmatrix}, \quad A\mathbf{b}_2 = \begin{bmatrix} 15 \\ -1 \end{bmatrix}, \quad A\mathbf{b}_3 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

and hence,

$$A(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = \left[\begin{array}{c|c|c} 6 & 15 & 5 \\ -2 & -1 & 1 \end{array}\right]$$

In general, if A is an $m \times n$ matrix and B is an $n \times r$ matrix that has been partitioned into columns $(\mathbf{b}_1, \ldots, \mathbf{b}_r)$, then the block multiplication of A times B is given by

$$\boxed{AB = (A\mathbf{b}_1, A\mathbf{b}_2, \ldots, A\mathbf{b}_r)}$$

In particular,

$$(\mathbf{a}_1, \ldots, \mathbf{a}_n) = A = AI = (A\mathbf{e}_1, \ldots, A\mathbf{e}_n)$$

Let A be an $m \times n$ matrix. If we partition A into rows, then

$$A = \begin{bmatrix} \vec{\mathbf{a}}_1 \\ \vec{\mathbf{a}}_2 \\ \vdots \\ \vec{\mathbf{a}}_m \end{bmatrix}$$

If B is an $n \times r$ matrix, the ith row of the product AB is determined by multiplying the ith row of A times B. Thus the ith row of AB is $\mathbf{a}(i,:)B$. In general, the product AB can be partitioned into rows as follows:

$$AB = \begin{bmatrix} \vec{\mathbf{a}}_1 B \\ \vec{\mathbf{a}}_2 B \\ \vdots \\ \vec{\mathbf{a}}_m B \end{bmatrix}$$

To illustrate this result, let us look at an example. If

$$A = \begin{bmatrix} 2 & 5 \\ 3 & 4 \\ 1 & 7 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 & 2 & -3 \\ -1 & 1 & 1 \end{bmatrix}$$

then

$$\vec{\mathbf{a}}_1 B = \begin{bmatrix} 1 & 9 & -1 \end{bmatrix}$$

$$\vec{\mathbf{a}}_2 B = \begin{bmatrix} 5 & 10 & -5 \end{bmatrix}$$

$$\vec{\mathbf{a}}_3 B = \begin{bmatrix} -4 & 9 & 4 \end{bmatrix}$$

These are the row vectors of the product AB:

$$AB = \begin{bmatrix} \vec{\mathbf{a}}_1 B \\ \vec{\mathbf{a}}_2 B \\ \vec{\mathbf{a}}_3 B \end{bmatrix} = \left[\begin{array}{rrr} 1 & 9 & -1 \\ \hline 5 & 10 & -5 \\ \hline -4 & 9 & 4 \end{array}\right]$$

Block Multiplication

Let A be an $m \times n$ matrix and B an $n \times r$ matrix. It is often useful to partition A and B and express the product in terms of the submatrices of A and B. Consider the following four cases:

Case 1. If $B = \begin{pmatrix} B_1 & B_2 \end{pmatrix}$, where B_1 is an $n \times t$ matrix and B_2 is an $n \times (r-t)$ matrix, then

$$\begin{aligned} AB &= A(\mathbf{b}_1, \ldots, \mathbf{b}_t, \mathbf{b}_{t+1}, \ldots, \mathbf{b}_r) \\ &= (A\mathbf{b}_1, \ldots, A\mathbf{b}_t, A\mathbf{b}_{t+1}, \ldots, A\mathbf{b}_r) \\ &= (A(\mathbf{b}_1, \ldots, \mathbf{b}_t), A(\mathbf{b}_{t+1}, \ldots, \mathbf{b}_r)) \\ &= \begin{pmatrix} AB_1 & AB_2 \end{pmatrix} \end{aligned}$$

Thus,

$$\boxed{A\begin{pmatrix} B_1 & B_2 \end{pmatrix} = \begin{pmatrix} AB_1 & AB_2 \end{pmatrix}}$$

Case 2. If $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$, where A_1 is a $k \times n$ matrix and A_2 is an $(m-k) \times n$ matrix, then

$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} B = \left[\begin{array}{c} \vec{\mathbf{a}}_1 \\ \vdots \\ \vec{\mathbf{a}}_k \\ \hline \vec{\mathbf{a}}_{k+1} \\ \vdots \\ \vec{\mathbf{a}}_m \end{array}\right] B = \left[\begin{array}{c} \vec{\mathbf{a}}_1 B \\ \vdots \\ \vec{\mathbf{a}}_k B \\ \hline \vec{\mathbf{a}}_{k+1} B \\ \vdots \\ \vec{\mathbf{a}}_m B \end{array}\right]$$

$$= \begin{bmatrix} \begin{bmatrix} \vec{\mathbf{a}}_1 \\ \vdots \\ \vec{\mathbf{a}}_k \end{bmatrix} B \\ \begin{bmatrix} \vec{\mathbf{a}}_{k+1} \\ \vdots \\ \vec{\mathbf{a}}_m \end{bmatrix} B \end{bmatrix} = \begin{bmatrix} A_1 B \\ A_2 B \end{bmatrix}$$

Thus,

$$\boxed{\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} B = \begin{bmatrix} A_1 B \\ A_2 B \end{bmatrix}}$$

Case 3. Let $A = \begin{pmatrix} A_1 & A_2 \end{pmatrix}$ and $B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$, where A_1 is an $m \times s$ matrix, A_2 is an $m \times (n - s)$ matrix, B_1 is an $s \times r$ matrix, and B_2 is an $(n - s) \times r$ matrix. If $C = AB$, then

$$c_{ij} = \sum_{l=1}^{n} a_{il} b_{lj} = \sum_{l=1}^{s} a_{il} b_{lj} + \sum_{l=s+1}^{n} a_{il} b_{lj}$$

Thus, c_{ij} is the sum of the (i, j) entry of $A_1 B_1$ and the (i, j) entry of $A_2 B_2$. Therefore,

$$AB = C = A_1 B_1 + A_2 B_2$$

and it follows that

$$\boxed{\begin{pmatrix} A_1 & A_2 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = A_1 B_1 + A_2 B_2}$$

Case 4. Let A and B both be partitioned as follows:

$$A = \begin{array}{cc} \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] & \begin{array}{l} k \\ m-k \end{array} \\ \begin{array}{cc} s & n-s \end{array} & \end{array}, \quad B = \begin{array}{cc} \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right] & \begin{array}{l} s \\ n-s \end{array} \\ \begin{array}{cc} t & r-t \end{array} & \end{array}$$

Let

$$A_1 = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \qquad A_2 = \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix}$$

$$B_1 = \begin{bmatrix} B_{11} & B_{12} \end{bmatrix} \qquad B_2 = \begin{bmatrix} B_{21} & B_{22} \end{bmatrix}$$

It follows from case 3 that

$$AB = \begin{bmatrix} A_1 & A_2 \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = A_1 B_1 + A_2 B_2$$

It follows from cases 1 and 2 that

$$A_1 B_1 = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} B_1 = \begin{bmatrix} A_{11} B_1 \\ A_{21} B_1 \end{bmatrix} = \begin{bmatrix} A_{11} B_{11} & A_{11} B_{12} \\ A_{21} B_{11} & A_{21} B_{12} \end{bmatrix}$$

$$A_2 B_2 = \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} B_2 = \begin{bmatrix} A_{12} B_2 \\ A_{22} B_2 \end{bmatrix} = \begin{bmatrix} A_{12} B_{21} & A_{12} B_{22} \\ A_{22} B_{21} & A_{22} B_{22} \end{bmatrix}$$

Therefore,

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}$$

In general, if the blocks have the proper dimensions, the block multiplication can be carried out in the same manner as ordinary matrix multiplication. That is, if

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1t} \\ \vdots & & \\ A_{s1} & \cdots & A_{st} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} B_{11} & \cdots & B_{1r} \\ \vdots & & \\ B_{t1} & \cdots & B_{tr} \end{pmatrix}$$

then

$$AB = \begin{bmatrix} C_{11} & \cdots & C_{1r} \\ \vdots & & \\ C_{s1} & \cdots & C_{sr} \end{bmatrix}$$

where

$$C_{ij} = \sum_{k=1}^{t} A_{ik} B_{kj}$$

The multiplication can be carried out in this manner only if the number of columns of A_{ik} equals the number of rows of B_{kj} for each k.

EXAMPLE 1 Let

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

and

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right]$$

Partition A into four blocks and perform the block multiplication.

Solution

Since each B_{kj} has two rows, the A_{ik}'s must have two columns. Thus, we have one of two possibilities:

$$\text{(i)} \quad \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ \hline 2 & 2 & 1 & 1 \\ 3 & 3 & 2 & 2 \end{array}\right]$$

in which case

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ \hline 2 & 2 & 1 & 1 \\ 3 & 3 & 2 & 2 \end{array}\right] \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right] = \left[\begin{array}{cc|cc} 8 & 6 & 4 & 5 \\ \hline 10 & 9 & 6 & 7 \\ 18 & 15 & 10 & 12 \end{array}\right]$$

or

$$\text{(ii)} \quad \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ \hline 3 & 3 & 2 & 2 \end{array}\right]$$

in which case

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ \hline 3 & 3 & 2 & 2 \end{array}\right] \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right] = \left[\begin{array}{cc|cc} 8 & 6 & 4 & 5 \\ 10 & 9 & 6 & 7 \\ \hline 18 & 15 & 10 & 12 \end{array}\right] \qquad \blacksquare$$

EXAMPLE 2 Let A be an $n \times n$ matrix of the form

$$\begin{pmatrix} A_{11} & O \\ O & A_{22} \end{pmatrix}$$

where A_{11} is a $k \times k$ matrix ($k < n$). Show that A is nonsingular if and only if A_{11} and A_{22} are nonsingular.

Solution

If A_{11} and A_{22} are nonsingular, then

$$\begin{pmatrix} A_{11}^{-1} & O \\ O & A_{22}^{-1} \end{pmatrix} \begin{pmatrix} A_{11} & O \\ O & A_{22} \end{pmatrix} = \begin{pmatrix} I_k & O \\ O & I_{n-k} \end{pmatrix} = I$$

and

$$\begin{pmatrix} A_{11} & O \\ O & A_{22} \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & O \\ O & A_{22}^{-1} \end{pmatrix} = \begin{pmatrix} I_k & O \\ O & I_{n-k} \end{pmatrix} = I$$

so A is nonsingular and

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} & O \\ O & A_{22}^{-1} \end{pmatrix}$$

Conversely, if A is nonsingular, then let $B = A^{-1}$ and partition B in the same manner as A. Since

$$BA = I = AB$$

it follows that

$$\begin{Bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{Bmatrix} \begin{Bmatrix} A_{11} & O \\ O & A_{22} \end{Bmatrix} = \begin{Bmatrix} I_k & O \\ O & I_{n-k} \end{Bmatrix} = \begin{Bmatrix} A_{11} & O \\ O & A_{22} \end{Bmatrix} \begin{Bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{Bmatrix}$$
$$\begin{Bmatrix} B_{11}A_{11} & B_{12}A_{22} \\ B_{21}A_{11} & B_{22}A_{22} \end{Bmatrix} = \begin{Bmatrix} I_k & O \\ O & I_{n-k} \end{Bmatrix} = \begin{Bmatrix} A_{11}B_{11} & A_{11}B_{12} \\ A_{22}B_{21} & A_{22}B_{22} \end{Bmatrix}$$

Thus,

$$\begin{aligned} B_{11}A_{11} &= I_k = A_{11}B_{11} \\ B_{22}A_{22} &= I_{n-k} = A_{22}B_{22} \end{aligned}$$

Hence, A_{11} and A_{22} are both nonsingular with inverses B_{11} and B_{22}, respectively. ■

Given two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, it is possible to perform a matrix multiplication of the vectors if we transpose one of the vectors first. The matrix product $\mathbf{x}^T\mathbf{y}$ is the product of a row vector (a $1 \times n$ matrix) and a column vector (an $n \times 1$ matrix). The result will be a 1×1 matrix, or simply a scalar:

$$\mathbf{x}^T\mathbf{y} = (x_1, x_2, \ldots, x_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

This type of product is referred to as a *scalar product* or an *inner product*. The scalar product is one of the most commonly performed operations. For example, when we multiply two matrices, each entry of the product is computed as a scalar product (a row vector times a column vector).

It is also useful to multiply a column vector times a row vector. The matrix product $\mathbf{x}\mathbf{y}^T$ is the product of an $n \times 1$ matrix times a $1 \times n$ matrix. The result is a full $n \times n$ matrix.

$$\mathbf{x}\mathbf{y}^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}(y_1, y_2, \ldots, y_n) = \begin{pmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & x_2y_2 & \cdots & x_2y_n \\ \vdots & & & \\ x_ny_1 & x_ny_2 & \cdots & x_ny_n \end{pmatrix}$$

The product $\mathbf{x}\mathbf{y}^T$ is referred to as the *outer product* of $\mathbf{x}$ and $\mathbf{y}$. The outer product matrix

$$XY^T = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)\begin{pmatrix} \mathbf{y}_1^T \\ \mathbf{y}_2^T \\ \vdots \\ \mathbf{y}_n^T \end{pmatrix} = \mathbf{x}_1\mathbf{y}_1^T + \mathbf{x}_2\mathbf{y}_2^T + \cdots + \mathbf{x}_n\mathbf{y}_n^T$$

EXAMPLE 3 Given

$$X = \begin{pmatrix} 3 & 1 \\ 2 & 4 \\ 1 & 2 \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 1 \end{pmatrix}$$

compute the outer product expansion of XY^T.

Solution

$$\begin{aligned} XY^T &= \begin{pmatrix} 3 & 1 \\ 2 & 4 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}\begin{pmatrix} 2 & 4 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 6 & 9 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{pmatrix} + \begin{pmatrix} 2 & 4 & 1 \\ 8 & 16 & 4 \\ 4 & 8 & 2 \end{pmatrix} \end{aligned}$$

■

(b) Let A and B be $n \times n$ matrices and define $2n \times 2n$ matrices S and M by

$$S = \begin{bmatrix} I & A \\ O & I \end{bmatrix}, \qquad M = \begin{bmatrix} AB & O \\ B & O \end{bmatrix}$$

Determine the block form of S^{-1} and use it to compute the block form of the product $S^{-1}MS$.

Ans. Let $S^{-1} = T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$. Then

$$\begin{aligned}
ST &= I = TS \\
\Rightarrow \quad \begin{bmatrix} I & A \\ O & I \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} &= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} I & A \\ O & I \end{bmatrix} \\
\Rightarrow \quad \begin{bmatrix} T_{11} + AT_{21} & T_{12} + AT_{22} \\ T_{21} & T_{22} \end{bmatrix} &= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} T_{11} & T_{11}A + T_{12} \\ T_{21} & T_{21}A + T_{22} \end{bmatrix} \\
\Rightarrow T_{11} + AT_{21} = I, T_{12} + AT_{22} = 0, T_{21} = 0, T_{22} = I, &\quad \text{and} \quad T_{11} = I, T_{11}A + T_{12} = 0, T_{21} = 0, T_{21}A + \\
T_{11} = I, T_{12} = -A, &\qquad T_{21} = 0, T_{22} = I.
\end{aligned}$$

Thus,

$$S^{-1} = \begin{bmatrix} I & -A \\ O & I \end{bmatrix}$$

It follows

$$\begin{aligned} S^{-1}MS &= \begin{bmatrix} I & -A \\ O & I \end{bmatrix} \begin{bmatrix} AB & O \\ B & O \end{bmatrix} \begin{bmatrix} I & A \\ O & I \end{bmatrix} \\ &= \begin{bmatrix} I & -A \\ O & I \end{bmatrix} \begin{bmatrix} AB & ABA \\ B & BA \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix} \end{aligned}$$

Do all the exercises of S. J. Leon Book Section 1.1 to 1.6

15. Let O be the $k \times k$ matrix whose entries are all 0, I be the $k \times k$ identity matrix, and B be a $k \times k$ matrix with the property that $B^2 = O$. If

$$A = \begin{pmatrix} O & I \\ I & B \end{pmatrix}$$

determine the block form of $A^{-1} + A^2 + A^3$.

16. Let A and B be $n \times n$ matrices and define $2n \times 2n$ matrices S and M by

$$S = \begin{pmatrix} I & A \\ O & I \end{pmatrix}, \qquad M = \begin{pmatrix} AB & O \\ B & O \end{pmatrix}$$

Determine the block form of S^{-1} and use it to compute the block form of the product $S^{-1}MS$.

17. Let

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

where A_{11} is a $k \times k$ nonsingular matrix. Show that A can be factored into a product of the form

$$\begin{pmatrix} I & O \\ B & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ O & C \end{pmatrix}$$

where

$$B = A_{21}A^{-1}{}_{11} \quad \text{and} \quad C = A_{22} - A_{21}A^{-1}{}_{11}A_{12}$$

19. Let A be an $n \times n$ matrix and $\mathbf{x} \in \mathbb{R}^n$.

(a) A scalar c can also be considered as a 1×1 matrix $C = (c)$, and a vector $\mathbf{b} \in \mathbb{R}^n$ can be considered as an $n \times 1$ matrix B. Although the matrix multiplication CB is not defined, show that the matrix product BC is equal to $c\mathbf{b}$, the scalar multiplication of c times $\mathbf{b}$.

(b) Partition A into columns and $\mathbf{x}$ into rows and perform the block multiplication of A times $\mathbf{x}$.

(c) Show that

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

20. If A is an $n \times n$ matrix with the property that $A\mathbf{x} = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^n$, show that $A = O$. [*Hint*: Let $\mathbf{x} = \mathbf{e}_j$ for $j = 1, \ldots, n$.]

21. Let B and C be $n \times n$ matrices with the property that $B\mathbf{x} = C\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. Show that $B = C$.

GROUP

Definition. A nonempty set G is said to be a *group* if in G there is defined an operation $*$ such that:

(a) $a, b \in G$ implies that $a * b \in G$. (We describe this by saying that G is *closed* under $*$.)

(b) Given $a, b, c \in G$, then $a * (b * c) = (a * b) * c$. (This is described by saying that the *associative law* holds in G.)

(c) There exists a special element $e \in G$ such that $a * e = e * a = a$ for all $a \in G$ (e is called the *identity* or *unit element* of G).

(d) For every $a \in G$ there exists an element $b \in G$ such that $a * b = b * a = e$. (We write this element b as a^{-1} and call it the *inverse* of a in G.)

**Ref. Book: *Abstract Algebra by I. N. Herstein*

Examples of Groups

1. Let $\mathbb{Z}$ be the set of all integers and let $*$ be the ordinary addition, $+$, in $\mathbb{Z}$. That $\mathbb{Z}$ is closed and associative under $*$ are basic properties of the integers. What serves as the unit element, e, of $\mathbb{Z}$ under $*$? Clearly, since $a = a * e = a + e$, we have $e = 0$, and 0 is the required identity element under addition. What about a^{-1}? Here too, since $e = 0 = a * a^{-1} = a + a^{-1}$, the a^{-1} in this instance is $-a$, and clearly $a * (-a) = a + (-a) = 0$.

2. Let $\mathbb{Q}$ be the set of all rational numbers and let the operation $*$ on $\mathbb{Q}$ be the ordinary addition of rational numbers. As above, $\mathbb{Q}$ is easily shown to be a group under $*$. Note that $\mathbb{Z} \subset \mathbb{Q}$ and both $\mathbb{Z}$ and $\mathbb{Q}$ are groups under the same operation $*$.

3. Let $\mathbb{Q}'$ be the set of all *nonzero* rational numbers and let the operation $*$ on $\mathbb{Q}'$ be the ordinary multiplication of rational numbers. By the familiar properties of the rational numbers we see that $\mathbb{Q}'$ forms a group relative to $*$.

Definition. A *field* is a set F, on which two operations $+$ and $\cdot$ (called *addition* and *multiplication*, respectively) are defined so that for each pair of elements x, y in F there are unique elements $x+y$ and $x \cdot y$ (often written xy) in F for which the following conditions hold for all elements x, y, z in F:

(i) $x+y=y+x$ (commutativity of addition)

(ii) $(x+y)+z=x+(y+z)$ (associativity of addition)

(iii) There is an element $0 \in F$, called zero, such that $x+0=x$. (existence of an additive identity)

(iv) For each x, there is an element $-x \in F$ such that $x+(-x)=0$. (existence of additive inverses)

(v) $xy=yx$ (commutativity of multiplication)

(vi) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ (associativity of multiplication)

(vii) $(x+y) \cdot z = x \cdot z + y \cdot z$ and $x \cdot (y+z) = x \cdot y + x \cdot z$ (distributivity)

(viii) There is an element $1 \in F$, such that $1 \neq 0$ and $x \cdot 1 = x$. (existence of a multiplicative identity)

(ix) If $x \neq 0$, then there is an element $x^{-1} \in F$ such that $x \cdot x^{-1} = 1$. (existence of multiplicative inverses)

Vector Space:

A *vector space* is a set V along with an addition on V and a scalar multiplication on V such that the following properties hold (over field $\mathbf{F}$).

closed under addition & scalar multiplication

- An ***addition*** on a set V is a function that assigns an element $u+v \in V$ to each pair of elements $u, v \in V$.
- A ***scalar multiplication*** on a set V is a function that assigns an element $\lambda v \in V$ to each $\lambda \in \mathbf{F}$ and each $v \in V$.

commutativity

$u + v = v + u$ for all $u, v \in V$.

associativity

$(u + v) + w = u + (v + w)$ and $(ab)v = a(bv)$ for all $u, v, w \in V$ and for all $a, b \in \mathbf{F}$.

additive identity

There exists an element $0 \in V$ such that $v + 0 = v$ for all $v \in V$.

additive inverse

For every $v \in V$, there exists $w \in V$ such that $v + w = 0$.

multiplicative identity

$1v = v$ for all $v \in V$.

distributive properties

$a(u + v) = au + av$ and $(a + b)v = av + bv$ for all $a, b \in \mathbf{F}$ and all $u, v \in V$.