MATH 106 LINEAR ALGEBRA

LECTURE NOTES

FALL 2010-2011[1]

[1]These Lecture Notes are not in a final form being still subject of improvement

Contents

Chapter 1

Systems of linear equations and matrices

1.1 Introduction to systems of linear equations

Definition 1.1.1 A *linear equation* in the n variables $x_1, \ldots, x_n$, equally called *unknowns*, is the problem of finding the values of $x_1, \ldots, x_n$ such that $a_1x_1 + \ldots + a_nx_n = b$, where $a_1, \ldots, a_n, b$ are constants. A *solution* of a linear equation $a_1x_1 + \ldots + a_nx_n = b$ is a sequence $t_1, \ldots, t_n$ of n real numbers such that the equation is satisfied when we replace $x_1 = t_1, \ldots, x_n = t_n$. The set of all solutions of the equation is called its *solution set* or *the general solution*.

Definition 1.1.2 A finite set of linear equations in the variables $x_1, \ldots, x_n$ is called a *system of linear equations* or a *a linear system*. A *solution* of a linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \vdots \qquad\qquad\qquad & \ \ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned} \tag{1.1}$$

with m linear equations and n unknowns, is a sequence $t_1, \ldots, t_n$ of n real numbers such that each equation of the system is satisfied when we replace $x_1 = t_1, \ldots, x_n = t_n$. The set of all solutions of the linear system is called its *solution set* or *the general solution*. A system of equations is said to be *consistent* if it has a solution at least. Otherwise it is called *inconsistent*. The augmented matrix of the linear system (1.3) is

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix}$$

Remark 1.1.3 The solution set of a linear system unchanges if we perform on the system one of the following operations:

1. Multiply an equation with a nonzero constant;
2. Interchange two equations;
3. Add a multiple of one equation to another.

The corresponding operations at the level of augmented matrix are:

1. Multiply a row with a nonzero constant;
2. Interchange two rows;
3. Add a multiple of one row to another.

Example 1.1.4 Solve the following linear system

$$\begin{array}{rcrcrcrcr} x & + & 2y & - & 3z & + & aw & = & 4 \\ 3x & - & y & + & 5z & + & 10aw & = & -2 \\ 4x & + & y & + & 2z & + & 11aw & = & 2 \end{array}.$$

Solution: The augmented matrix of the system is $\left[\begin{array}{rrrr|r} 1 & 2 & -3 & a & 4 \\ 3 & -1 & 5 & 10a & -2 \\ 4 & 1 & 2 & 11a & 2 \end{array}\right]$ and we have successively:

$$
\left[\begin{array}{cccc|c} 1 & 2 & -3 & a & 4 \\ 3 & -1 & 5 & 10a & -2 \\ 4 & 1 & 2 & 11a & 2 \end{array}\right]
\qquad
\begin{array}{rcrcrcrcr} x & + & 2y & - & 3z & + & aw & = & 4 \\ 3x & - & y & + & 5z & + & 10aw & = & -2 \\ 4x & + & y & + & 2z & + & 11aw & = & 2 \end{array}
$$

$$-3r_1 + r_2 \to r_2 \downarrow -4r_1 + r_3 \to r_3 \qquad -3e_1 + e_2 \to e_2 \downarrow -4e_1 + e_3 \to e_3$$

$$\downarrow \qquad \downarrow$$

$$
\left[\begin{array}{cccc|c} 1 & 2 & -3 & a & 4 \\ 0 & -7 & 14 & 7a & -14 \\ 0 & -7 & 14 & 7a & -14 \end{array}\right]
\qquad
\begin{array}{rcrcrcrcr} x & + & 2y & - & 3z & + & aw & = & 4 \\ & - & 7y & + & 14z & + & 7aw & = & -14 \\ & - & 7y & + & 14z & + & 7aw & = & -14 \end{array}
$$

$$-r_2 + r_3 \to r_3 \downarrow \qquad \downarrow -e_2 + e_3 \to e_3$$

$$\downarrow \qquad \downarrow$$

$$
\left[\begin{array}{cccc|c} 1 & 2 & -3 & a & 4 \\ 0 & -7 & 14 & 7a & -14 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]
\qquad
\begin{array}{rcrcrcrcr} x & + & 2y & - & 3z & + & aw & = & 4 \\ & - & 7y & + & 14z & + & 7aw & = & -14 \\ & & & & & & 0 & = & 0 \end{array}
$$

$$-\tfrac{1}{7}r_2 \downarrow \qquad \downarrow -\tfrac{1}{7}e_2$$

$$\downarrow \qquad \downarrow$$

$$
\left[\begin{array}{cccc|c} 1 & 2 & -3 & a & 4 \\ 0 & 1 & -2 & -a & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]
\qquad
\begin{array}{rcrcrcrcr} x & + & 2y & - & 3z & + & aw & = & 4 \\ & & y & - & 2z & - & aw & = & 2 \\ & & & & & & 0 & = & 0 \end{array}
$$

$$-2r_2 + r_1 \to r_1 \downarrow \qquad \downarrow -2e_2 + e_1 \to e_1$$

$$\downarrow \qquad \downarrow$$

$$
\left[\begin{array}{cccc|c} 1 & 0 & 1 & 3a & 0 \\ 0 & 1 & -2 & -a & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]
\qquad
\begin{array}{rcrcrcrcr} x & & & + & z & + & 3aw & = & 0 \\ & & y & - & 2z & - & aw & = & 2 \\ & & & & & & 0 & = & 0. \end{array}
$$

Thus, the equivalent system is $\begin{array}{llll} x & +z & +3aw & = 0 \\ y & -2z & -aw & = 2 \end{array}$ and its solutions are

$$x = -t - 3as,\ y = 2 + 2t + as,\ z = t,\ w = s \text{ for } s, t \in \mathbf{R}.$$

1.2 Gaussian elimination

Definition 1.2.1 A matrix is said to be in *reduced row-echelon form* if it has the following properties:

1. If a row does not consist entirely of zeros, then the first nonzero number in the row is a 1, called a *leading* 1.

2. If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.

3. In any two successive rows that do not consist entirely of zeros, the leading 1 in the lower row occurs farther to the right than the leading 1 in the higher row.

4. Each column that contains a leading 1 has zero everywhere else.

A matrix satisfying just the first 3 properties, namely 1,2 and 3, is said to be in *row-echelon form*

We are now going to provide a number of steps in order to reduce a matrix to a (reduced) row-echelon form. These are:

1. Locate the lefmost column that does not consist entirely of zeros;

2. Interchange the top row with another row, if necessary, to bring a nonzero entry to the top of the column found in Step 1;

3. If the entry that is now at the top of the column found in Step 1 is a, multiply the first row by $\frac{1}{a}$ in order to introduce a leading 1;

4. Add suitable multiples of the top row to the rows below so that all entries below the leading 1 become zeros;

5. Cover the top row in the matrix and begin again with the Step 1 applied to the submatrix that remains. Continue in this way until the entire matrix is in row-echelon form.

6. Once you got the matrix in row-echelon form, beginning with the last nonzero row and working upward, add suitable multiples of each row to the rows above to introduce zeros above the leading 1's.

The variables corresponding to the leading 1's in row echelon form are called *leading variables* and the others are called *free variables.*

The above procedure for reducing a matrix to reduced row-echelon form is called *Gauss-Jordan elimination.* If we use only the first five steps, the procedure produces a row-echelon form and is called *Gaussian elimination.*

Examples 1.2.2 1. Solve the linear system

$$\begin{array}{rcrcrcrcr} x & - & y & + & 2z & - & w & = & -1 \\ 2x & + & y & - & 2z & - & 2w & = & -2 \\ -x & + & 2y & - & 4z & + & w & = & 1 \\ 3x & & & & & - & 3w & = & -3 \end{array} \tag{1.2}$$

Solution: The augmented matrix is $\left[\begin{array}{cccc|c} 1 & -1 & 2 & -1 & -1 \\ 2 & 1 & -2 & -2 & -2 \\ -1 & 2 & -4 & 1 & 1 \\ 3 & 0 & 0 & -3 & -3 \end{array}\right]$ and it can be reduced at reduced row-echelon form as follows:

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & -1 & -1 \\ 2 & 1 & -2 & -2 & -2 \\ -1 & 2 & -4 & 1 & 1 \\ 3 & 0 & 0 & -3 & -3 \end{array}\right] \xrightarrow[\substack{r_1 + r_3 \to r_3 \\ -3r_1 + r_4 \to r_4}]{-2r_1 + r_2 \to r_2} \left[\begin{array}{cccc|c} 1 & -1 & 2 & -1 & -1 \\ 0 & 3 & -6 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 3 & -6 & 0 & 0 \end{array}\right] \xrightarrow{\frac{1}{3}r_2} \left[\begin{array}{cccc|c} 1 & -1 & 2 & -1 & -1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 3 & -6 & 0 & 0 \end{array}\right]$$

$$\xrightarrow[-3r_2 + r_4 \to r_4]{-r_2 + r_3 \to r_3} \left[\begin{array}{cccc|c} 1 & -1 & 2 & -1 & -1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{r_2 + r_1 \to r_1} \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

such that the corresponding linear system, equivalent with the initial one, is

$$\begin{array}{ccccccc} x & & & - & w & = & -1 \\ & y & - & 2z & & = & 0 \end{array} \quad \text{and its solutions are} \quad \begin{array}{l} x = t - 1 \\ y = 2s \\ z = s \\ w = t \end{array}, s, t \in \mathbb{R},$$

being infinitely many in this case. It is sometimes preferable to solve a linear system by using Gauss elimination procedure for the augmented matrix to bring it just in a row-echelon form and to use the so called *back-substitution method* afterward. For the linear system 1.2, a row-echelon form of the augmented matrix is

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & -1 & -1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

and its corresponding linear system is

$$\begin{array}{ccccccccc} x & - & y & + & 2z & - & w & = & -1 \\ & & y & - & 2z & & & = & 0 \end{array}$$

The back-substitution method consists in the following three steps:

(a) Solve the equations for the leading variables in terms of the free variables;

(b) Beginning with the bottom equation and working upward, successively substitute each equation into all the equations above it.

(c) Assign arbitrary values to the free variables, if any.

In our particular case x, y are the leading variables and z, w are the free variables. Therefore the step (1a) is

$$\begin{aligned} x &= y - 2z + w - 1 \\ y &= 2z \end{aligned}$$

The step (1b) is

$$\begin{aligned} x &= 2z - 2z + w - 1 \\ y &= 2z \end{aligned} \quad \text{or equivalently} \quad \begin{aligned} x &= w - 1 \\ y &= 2z \end{aligned}$$

Finally, the step (1c) consists in assigning to the free variables z, w the arbitrary values $z = s$ and $w = t$ such that the solutions of the initial linear system are:

$$\begin{aligned} &x = t - 1 \\ &y = 2s \\ &z = s \\ &w = t \end{aligned} \quad , s, t \in \mathbb{R}.$$

2. Solve the linear system

$$\begin{array}{rcrcrl} -x & + & 7y & - & 2z & = 1 \\ 3x & - & y & + & z & = 4 \\ 2x & + & 6y & - & z & = 5 \end{array}$$

Solution: The augmented matrix is $\left[\begin{array}{rrr|r} -1 & 7 & -2 & 1 \\ 3 & -1 & 1 & 4 \\ 2 & 6 & -1 & 5 \end{array}\right]$ and it can be reduced at reduced row-echelon form as follows:

$$\left[\begin{array}{rrr|r} -1 & 7 & -2 & 1 \\ 3 & -1 & 1 & 4 \\ 2 & 6 & -1 & 5 \end{array}\right] \xrightarrow{-r_1} \left[\begin{array}{rrr|r} 1 & -7 & 2 & -1 \\ 3 & -1 & 1 & 4 \\ 2 & 6 & -1 & 5 \end{array}\right] \xrightarrow[-2r_1 + r_3 \to r_3]{-3r_1 + r_2 \to r_2} \left[\begin{array}{rrr|r} 1 & -7 & 2 & -1 \\ 0 & 20 & -5 & 7 \\ 0 & 20 & -5 & 7 \end{array}\right] \xrightarrow{-r_2 + r_3 \to r_3}$$

$$\left[\begin{array}{rrr|r} 1 & -7 & 2 & -1 \\ 0 & 20 & -5 & 7 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{1/20 r_2} \left[\begin{array}{rrr|r} 1 & -7 & 2 & -1 \\ 0 & 1 & -\frac{1}{4} & \frac{7}{20} \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{7r_2 + r_1 \to r_1} \left[\begin{array}{rrr|r} 1 & 0 & \frac{1}{4} & \frac{29}{20} \\ 0 & 1 & -\frac{1}{4} & 7/20 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

such that the corresponding linear system, equivalent with the initial one, is

$$\begin{array}{rcrcr} x & & + \frac{1}{4}z & = & \frac{29}{20} \\ & y & - \frac{1}{4}z & = & 7/20 \end{array} \quad \text{and its solutions are} \quad \begin{aligned} &x = \tfrac{29-5t}{20} \\ &y = \tfrac{7+5t}{20} \\ &z = t. \end{aligned} \quad t \in \mathbb{R}.$$

3. Solve the linear system

$$\begin{array}{rcrcrcrl} x_1 & - & 2x_2 & + & x_3 & - & 4x_4 & = 1 \\ x_1 & + & 3x_2 & + & 7x_3 & + & 2x_4 & = 2 \\ x_1 & - & 12x_2 & - & 11x_3 & - & 16x_4 & = 5 \end{array}$$

Solution: The augmented matrix is $\left[\begin{array}{rrrr|r} 1 & -2 & 1 & -4 & 1 \\ 1 & 3 & 7 & 2 & 2 \\ 1 & -12 & -11 & -16 & 5 \end{array}\right]$ and it can be reduced at reduced row-echelon form as follows:

$$\left[\begin{array}{rrrr|r} 1 & -2 & 1 & -4 & 1 \\ 1 & 3 & 7 & 2 & 2 \\ 1 & -12 & -11 & -16 & 5 \end{array}\right] \xrightarrow[-r_1 + r_3 \to r_3]{-r_1 + r_2 \to r_2} \left[\begin{array}{rrrr|r} 1 & -2 & 1 & -4 & 1 \\ 0 & 5 & 6 & 6 & 1 \\ 0 & -10 & -12 & -12 & 4 \end{array}\right] \xrightarrow{2r_2 + r_3 \to r_3}$$

$$\xrightarrow{2r_2 + r_3 \to r_3} \left[\begin{array}{rrrr|r} 1 & -2 & 1 & -4 & 1 \\ 0 & 5 & 6 & 6 & 1 \\ 0 & 0 & 0 & 0 & 6 \end{array}\right].$$

Thus the system is inconsistent, that is it has no solutions at all, since $0 \neq 6$.

4. For which values of $a \in \mathbb{R}$ the following linear system has no solution ? Exactly one solution ? Infinitely many solutions ? When the system is consistent, find its solution set.

$$\begin{array}{rcrcrcr} x & - & y & - & 2z & = & 0 \\ 2x & - & 4y & - & 6z & = & -6 \\ 3x & - & 5y & + & (a^2 - 14)z & = & a - 4 \end{array}$$

Solution: The augmented matrix is $\left[\begin{array}{rrc|c} 1 & -1 & -2 & 0 \\ 2 & -4 & -6 & -6 \\ 3 & -5 & a^2 - 14 & a - 4 \end{array}\right]$ and it will be reduced by performing successively the following row-operations:

$$\left[\begin{array}{rrc|c} 1 & -1 & -2 & 0 \\ 2 & -4 & -8 & -6 \\ 3 & -5 & a^2 - 14 & a - 4 \end{array}\right] \xrightarrow[-3r_1 + r_3 \to r_3]{-2r_1 + r_2 \to r_2} \left[\begin{array}{rrc|c} 1 & -1 & -2 & 0 \\ 0 & -2 & -4 & -6 \\ 0 & -2 & a^2 - 8 & a - 4 \end{array}\right] \xrightarrow{-\frac{1}{2}r_2}$$

$$\xrightarrow{-\frac{1}{2}r_2} \left[\begin{array}{rrc|c} 1 & -1 & -2 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & -2 & a^2 - 8 & a - 4 \end{array}\right] \xrightarrow{2r_2 + r_3 \to r_3} \underbrace{\left[\begin{array}{rrc|c} 1 & -1 & -2 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & a^2 - 4 & a + 2 \end{array}\right]}_{(*)} \xrightarrow[\text{if } a \notin \{\pm 2\}]{\frac{1}{a^2 - 4}}$$

$$\xrightarrow[\text{if } a\notin\{\pm 2\}]{\frac{1}{a^2-4}} \left[\begin{array}{ccc|c} 1 & -1 & -2 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & \frac{1}{a-2} \end{array}\right] \xrightarrow[\text{if } a\notin\{\pm 2\}]{\substack{-2r_3+r_2\to r_2 \\ 2r_3+r_1\to r_1}} \left[\begin{array}{ccc|c} 1 & -1 & 0 & \frac{2}{a-2} \\ 0 & 1 & 0 & 3-\frac{2}{a-2} \\ 0 & 0 & 1 & \frac{1}{a-2} \end{array}\right] \xrightarrow[\text{if } a\notin\{\pm 2\}]{r_2+r_1\to r_1}$$

$$\xrightarrow[\text{if } a\notin\{\pm 2\}]{r_2+r_1\to r_1} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & \frac{3a-8}{a-2} \\ 0 & 0 & 1 & \frac{1}{a-2} \end{array}\right].$$

Therefore for $a \notin \{\pm 2\}$ the given linear system has the unique solutions $x = 3, y = \frac{3a-8}{a-2}, z = \frac{1}{a-2}$.

If $a = -2$, the $(*)$ reduced matrix becomes

$$\left[\begin{array}{ccc|c} 1 & -1 & -2 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{r_2+r_1\to r_1} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

and in this case the given linear system has infinitely many solutions

$$x = 3, y = 3 - 2s, z = s, s \in \mathbb{R}.$$

Finally, when $a = 2$, the $(*)$ reduced matrix becomes $\left[\begin{array}{ccc|c} 1 & -1 & -2 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 4 \end{array}\right]$ and linear system has no solutions at all in this case.

1.3 Matrices and matrix operations

Definition 1.3.1 A *matrix*

$$A=\begin{array}{cl} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} & \begin{array}{l} \leftarrow \textit{row } 1 \\ \leftarrow \textit{row } 2 \\ \quad\vdots \\ \leftarrow \textit{row } m \end{array} \\ \begin{array}{cccc} \uparrow & \uparrow & \cdots & \uparrow \\ \textit{col.}1 & \textit{col.}2 & \cdots & \textit{col.}n \end{array} & \end{array}$$

equally written $[a_{ij}]_{m\times n}$ (or simply $[a_{ij}]$), is a rectangular array of numbers. The numbers in the array are called *entries*. The entry in row i and column j of a matrix A is commonly denoted by $(A)_{ij}$, in our particular case $(A)_{ij} = a_{ij}$. The *size* of a matrix is described in terms of the number of rows and the number of columns it has. The above matrix A has size $m \times n$. A matrix

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

with only one column is called a *column matrix* (or a *column vector*), and a matrix

$$\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix}$$

with only one row is called *row matrix* (or a *row vector*). A matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix},$$

with n rows and n columns, that is of size $n \times n$, is called a *square matrix of order* n and the entries $a_{11}, a_{22}, \dots, a_{nn}$ are said to be on the *main diagonal* of A.

Definition 1.3.2 Two matrices are said to be *equal* if they have the same size and their corresponding entries are equal.

Definition 1.3.3 If

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \ B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix},$$

are matrices of the same size, then their *sum* is the matrix

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix},$$

and their *difference* is the matrix

$$A - B = \begin{bmatrix} a_{11} - b_{11} & a_{12} - b_{12} & \dots & a_{1n} - b_{1n} \\ a_{21} - b_{21} & a_{22} - b_{22} & \dots & a_{2n} - b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} - b_{m1} & a_{m2} - b_{m2} & \dots & a_{mn} - b_{mn} \end{bmatrix}.$$

One can shortly define the entries of the sum and the difference matrices in the following way:

$$(A + B)_{ij} = (A)_{ij} + (B)_{ij} \text{ and } (A - B)_{ij} = (A)_{ij} - (B)_{ij}.$$

If the matrices A and B have different sizes, then their sum and difference is **not** defined. If c is any scalar (number), then the *product* cA is the matrix

$$\begin{bmatrix} ca_{11} & ca_{12} & \dots & ca_{1n} \\ ca_{21} & ca_{22} & \dots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \dots & ca_{mn} \end{bmatrix}$$

obtained by multiplying each entry of A with c and its entries can be shortly written as $(cA)_{ij} = c(A)_{ij}$. If $A_1, A_2, \dots, A_n$ are matrices of the same size and $c_1, c_2, \dots, c_n$ are scalars, then the matrix $c_1A_1 + c_2A_2 + \dots + c_nA_n$ is defined and it is called a *linear combination* of $A_1, A_2, \dots, A_n$ with coefficients $c_1, c_2, \dots, c_n$.

Examples 1.3.4 Consider the matrices

$$A = \begin{bmatrix} 1 & -3 & 0 \\ 2 & 0 & 2 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 2 & -2 \\ -2 & 4 & 3 \end{bmatrix}, \ C = \begin{bmatrix} 1 & 2 & -3 \\ -1 & 0 & 4 \\ 2 & 5 & -1 \end{bmatrix}, \ D = \begin{bmatrix} -1 & 6 & 2 \\ 2 & 1 & 3 \\ 1 & 2 & -4 \end{bmatrix}.$$

1. *Compute, when is possible,* $A - B, C + D$*,* $B - C$*,* $A + C$*,*

   Solution:

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

$$D + E = \begin{bmatrix} 1 & 2 & -3 \\ -1 & 0 & 4 \\ 2 & 5 & -1 \end{bmatrix} + \begin{bmatrix} -1 & 6 & 2 \\ 2 & 1 & 3 \\ 1 & 2 & -4 \end{bmatrix} = \begin{bmatrix} 0 & 8 & -1 \\ 1 & 1 & 7 \\ 3 & 7 & -5 \end{bmatrix}.$$

   $B - C$ *and* $A + C$ *are not defined.*

2. *Compute, when possible,* $\frac{1}{3}A, (-1)B, 2A - B, 3C + D, 5B + \frac{1}{2006}D$*.* Solution:

$$\frac{1}{3}A = \begin{bmatrix} \frac{1}{3} & -1 & 0 \\ \frac{2}{3} & 0 & \frac{2}{3} \end{bmatrix}, (-1)B = \begin{bmatrix} -1 & -2 & 2 \\ 2 & -4 & -3 \end{bmatrix}$$

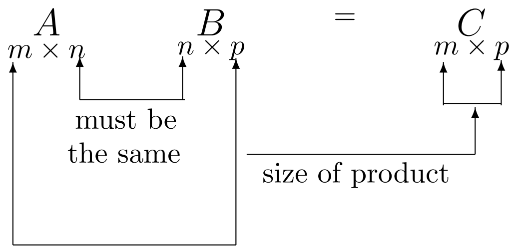
$$2A-B = 2\begin{bmatrix} 1 & -3 & 0 \\ 2 & 0 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 2 & -2 \\ -2 & 4 & 3 \end{bmatrix} = \begin{bmatrix} 2 & -6 & 0 \\ 4 & 0 & 4 \end{bmatrix} - \begin{bmatrix} 1 & 2 & -2 \\ -2 & 4 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -8 & 2 \\ 6 & -4 & 1 \end{bmatrix}.$$

$$3C+D = 3\begin{bmatrix} 1 & 2 & -3 \\ -1 & 0 & 4 \\ 2 & 5 & -1 \end{bmatrix} + \begin{bmatrix} -1 & 6 & 2 \\ 2 & 1 & 3 \\ 1 & 2 & -4 \end{bmatrix} = \begin{bmatrix} 3 & 6 & -9 \\ -3 & 0 & 12 \\ 6 & 15 & -3 \end{bmatrix} + \begin{bmatrix} -1 & 6 & 2 \\ 2 & 1 & 3 \\ 1 & 2 & -4 \end{bmatrix} = \begin{bmatrix} 2 & 12 & -7 \\ -1 & 1 & 15 \\ 7 & 17 & -7 \end{bmatrix}.$$

   $5B + \frac{1}{2006}D$ *is not defined.*

Definition 1.3.5 Lat A be an $m \times n$ matrix and B be an $n \times p$ matrix. Then the product AB is the $m \times p$ matrix C whose entry c_{ij} in row i and column j is obtained as follows: Sum the products formed by multiplying, in order, each entry in row i of the matrix A with the "corresponding" entry in column j of the matrix B. More precisely

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$



$$AB = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1j} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2j} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nj} & \dots & b_{np} \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{k=1}^{n} a_{1k}b_{k1} & \sum_{k=1}^{n} a_{1k}b_{k2} & \dots & \sum_{k=1}^{n} a_{1k}b_{kj} & \dots & \sum_{k=1}^{n} a_{1k}b_{kp} \\ \sum_{k=1}^{n} a_{2k}b_{k1} & \sum_{k=1}^{n} a_{2k}b_{k2} & \dots & \sum_{k=1}^{n} a_{2k}b_{kj} & \dots & \sum_{k=1}^{n} a_{2k}b_{kp} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{n} a_{ik}b_{k1} & \sum_{k=1}^{n} a_{ik}b_{k2} & \dots & \sum_{k=1}^{n} a_{ik}b_{kj} & \dots & \sum_{k=1}^{n} a_{ik}b_{kp} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{n} a_{mk}b_{k1} & \sum_{k=1}^{n} a_{mk}b_{k2} & \dots & \sum_{k=1}^{n} a_{mk}b_{kj} & \dots & \sum_{k=1}^{n} a_{mk}b_{kp} \end{bmatrix}.$$

Definition 1.3.6 The *transpose* of an $m \times n$ matrix A, denoted by A^T, is the $n \times m$ matrix whose $i - th$ row is the $i - th$ column of A. Thus $(A^T)_{ij} = (A)_{ji}$. Observe that for any matrix A we have $(A^T)^T = A$.

Example 1.3.7 If $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 & 0 & 1 \\ 3 & -1 & -1 & 0 \\ 4 & 3 & 2 & 1 \end{bmatrix}$, find, when possible, the matrices $AB, BA, (AB)^T, (BA)^T, A^TB^T, B^TA^T$.

Solution: then we have

$$AB = \underbrace{\begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 3 \end{bmatrix}}_{2\times 3} \underbrace{\begin{bmatrix} 2 & 1 & 0 & 1 \\ 3 & -1 & -1 & 0 \\ 4 & 3 & 2 & 1 \end{bmatrix}}_{3\times 4} =$$

$$= \begin{bmatrix} 1\cdot 2+0\cdot 3+(-1)\cdot 4 & 1\cdot 1+0\cdot(-1)+(-1)\cdot 3 & 1\cdot 0+0\cdot(-1)+(-1)\cdot 2 & 1\cdot 1+0\cdot 0+(-1)\cdot 1 \\ 0\cdot 2+2\cdot 3+3\cdot 4 & 0\cdot 1+2\cdot(-1)+3\cdot 3 & 0\cdot 0+2\cdot(-1)+3\cdot 2 & 0\cdot 1+2\cdot 0+3\cdot 1 \end{bmatrix} =$$

$$= \underbrace{\begin{bmatrix} -2 & -2 & -2 & 0 \\ 18 & 7 & 4 & 3 \end{bmatrix}}_{2\times 4}, \text{ such that } (AB)^T = \underbrace{\begin{bmatrix} -2 & 18 \\ -2 & 7 \\ -2 & 4 \\ 0 & 3 \end{bmatrix}}_{4\times 2}.$$

Observe that the product $BA, (BA)^T, A^T B^T$ are not defined. However, observe that

$$B^T = \underbrace{\begin{bmatrix} 2 & 3 & 4 \\ 1 & -1 & 3 \\ 0 & -1 & 2 \\ 1 & 0 & 1 \end{bmatrix}}_{4\times 3} \text{ and } A^T = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 2 \\ -1 & 3 \end{bmatrix}}_{3\times 2},$$

such that the product $B^T A^T$ is also defined and

$$B^T A^T = \underbrace{\begin{bmatrix} 2 & 3 & 4 \\ 1 & -1 & 3 \\ 0 & -1 & 2 \\ 1 & 0 & 1 \end{bmatrix}}_{4\times 3} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 2 \\ -1 & 3 \end{bmatrix}}_{3\times 2} = \underbrace{\begin{bmatrix} -2 & 18 \\ -2 & 7 \\ -2 & 4 \\ 0 & 3 \end{bmatrix}}_{4\times 2} = (AB)^T.$$

Definition 1.3.8 If A is a square matrix, then the *trace of* A, denoted by $tr(A)$, is defined to be the sum of the entries on the main diagonal, namely $tr(A) := (A)_{11} + (A)_{22} + \cdots + (A)_{nn}$, where n is the order of A. The trace of A is not defined if A is not a square matrix.

Example 1.3.9 Find, when possible, $tr(AA^T)$ and $tr(AB)$, where

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

We first observe that $AA^T = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ -1 & 3 \end{bmatrix} = \underbrace{\begin{bmatrix} 2 & -3 \\ -3 & 13 \end{bmatrix}}_{2\times 2}$, such that $tr(AA^T) = 2 + 13 = 15$. Since $AB = \underbrace{\begin{bmatrix} -2 & -2 & -2 & 0 \\ 18 & 7 & 4 & 3 \end{bmatrix}}_{2\times 4}$ is not a square matrix, its trace is not defined.

If $A_{ij}, 1 \leq i \leq m, 1 \leq j \leq n$ are matrices of suitable sizes, then we can form a new matrix

$$A=\left[\begin{array}{c:c:c:c} A_{11} & A_{12} & \cdots & A_{1n} \\ \hdashline A_{21} & A_{22} & \cdots & A_{2n} \\ \hdashline \vdots & \vdots & \ddots & \vdots \\ \hdashline A_{m1} & A_{m2} & \cdots & A_{mn} \end{array}\right]$$

and call *partition* of the matrix A the above kind of subdivision. For example

$$\left[\begin{array}{cc:cc:cc} -1 & 0 & 2 & 3 & -2 & 0 \\ 1 & 0 & 0 & -1 & 3 & 5 \\ \hdashline -1 & 1 & 0 & 1 & -1 & 0 \\ 0 & 5 & 2 & -7 & 6 & 0 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

where

$$A_{11} = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}, A_{12} = \begin{bmatrix} 2 & 3 \\ 0 & -1 \end{bmatrix}, A_{13} = \begin{bmatrix} -2 & 0 \\ 3 & 5 \end{bmatrix},$$

$$A_{21} = \begin{bmatrix} -1 & 1 \\ 0 & 5 \end{bmatrix}, A_{22} = \begin{bmatrix} 0 & 1 \\ 2 & -7 \end{bmatrix}, A_{23} = \begin{bmatrix} -1 & 0 \\ 6 & 0 \end{bmatrix}.$$

An application of matrix multiplication operation consists in transforming a system of linear equations into a matrix equation. Indeed, for a linear system we have successively:

$$\begin{array}{ll} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = b_2 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = b_m \end{array} \Leftrightarrow \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow \underbrace{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{X} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}}_{B} \Leftrightarrow AX = B.$$

The matrix

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

is called the *coefficient matrix* of the linear system. Observe that the augmented matrix of the system is $\left[\ A \vdots B\ \right]$.

1.4 Inverses. Rules of matrix arithmetic

Theorem 1.4.1 *Assuming that the sizes of matrices are such that the indicated operations can be performed, the following rules of matrix arithmetic are valid:*

1. $A + B = B + A$;
2. $(A + B) + C = A + (B + C)$;
3. $(AB)C = A(BC)$;
4. $A(B + C) = AB + AC$;
5. $(A + B)C = AC + BC$;
6. $A(B - C) = AB - AC$;
7. $(A - B)C = AC - BC$;
8. $a(B + C) = aB + aC$;
9. $a(B - C) = aB - aC$;
10. $(a + b)C = aC + bC$;
11. $(a - b)C = aC - bC$;
12. $a(bC) = (ab)C$;

13. $a(BC) = (aB)C = B(aC)$.

Remark 1.4.2 The matrix multiplication is not commutative. Indeed, for example if

$$A = \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix},$$

then

$$AB = \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 11 & 4 \end{bmatrix} \neq \begin{bmatrix} 3 & 6 \\ -3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix} = BA.$$

Define the $m \times n$ *zero* matrix to be the matrix $O_{m\times n} := \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}_{m\times n}$ and it will be equally denoted by O when the size is understood from context.

Theorem 1.4.3 *Assuming that the sizes of the matrices are such that the indicated operations can be performed, the following rules of matrix arithmetic are valid*

1. $A + O = O + A = A$*;*

2. $A - A = O$*;*

3. $O - A = -A$*;*

4. $AO = O$*.*

The *identity* matrix of order n is defined to be the square the $n \times n$ matrix

$$I_n := \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{n\times n},$$

and it will be equally denoted by I when the order is understood from context.

Remark 1.4.4 If A is an $m \times n$ matrix, then $AI_n = I_m A = A$.

Theorem 1.4.5 *If R is the reduced row-echelon form of an $n \times n$ matrix A, then either R has a row of zeros, or R is the identity matrix.*

Definition 1.4.6 A square matrix A is said to be *invertible* if there is another square matrix B, of the same size, such that $AB = BA = I$.

Remark 1.4.7 If B, C are both inverses of a A, then $B = C$. Indeed, we have successively: $B = BI = B(AC) = (BA)C = IC = C$. Therefore, the inverse of an invertible matrix A is unique and denoted by A^{-1}, being characterized by the equalities: $AA^{-1} = A^{-1}A = I$.

Example 1.4.8 If $a, b, c, d \in \mathbb{R}$ are such that $ad - bc \neq 0$, then the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible and $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}$. Indeed we have

$$\underbrace{\begin{bmatrix} a & b \\ c & d \end{bmatrix}}_{A}\begin{bmatrix} \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} = \begin{bmatrix} \frac{ad-bc}{ad-bc} & \frac{-ab+ba}{ad-bc} \\ \frac{cd-dc}{ad-bc} & \frac{-cb+da}{ad-bc} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

and similarly

$$\begin{bmatrix} \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}\underbrace{\begin{bmatrix} a & b \\ c & d \end{bmatrix}}_{A} = \begin{bmatrix} \frac{da-bc}{ad-bc} & \frac{db-bd}{ad-bc} \\ \frac{-ca+ac}{ad-bc} & \frac{-cb+ad}{ad-bc} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

If A is a square matrix, then its positive powers are defined to be $A^0 = I, A^n = \underbrace{AA\cdots A}_{n \text{ times}}$ for $n > 0$. Moreover, if A is invertible, then $A^{-n} = \underbrace{A^{-1}A^{-1}\cdots A^{-1}}_{n \text{ times}}$, also for $n > 0$.

Theorem 1.4.9 *If A is an invertible matrix, then*

1. *A^{-1} is also invertible and $(A^{-1})^{-1} = A$;*
2. *A^n is invertible and $(A^n)^{-1} = (A^{-1})^n$ for $n \in \{0, 1, 2, \ldots\}$;*
3. *A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$;*
4. *For any non-zero scalar k, the matrix kA is invertible and $(kA)^{-1} = \frac{1}{k}A^{-1}$;*
5. *$A^m A^n = A^{m+n}$ for any integers m, n;*
6. *$(A^m)^n = A^{mn}$ for any integers m, n.*

Theorem 1.4.10 *If the sizes of the involved matrices are such that the stated operations can be performed, then*

1. $(A^T)^T = A$;

2. $(A+B)^T = A^T + B^T$ *and* $(A-B)^T = A^T - B^T$;

3. $(kA)^T = kA^T$, *where* k *is any scalar;*

4. $(AB)^T = B^T A^T$.

Example 1.4.11 1. Find the matrix A knowing that $(I_2 + 2A)^{-1} = \begin{bmatrix} -1 & 2 \\ 4 & 5 \end{bmatrix}$.

Solution: Since $(I_2 + 2A)^{-1} = \begin{bmatrix} -1 & 2 \\ 4 & 5 \end{bmatrix}$, it follows that $[(I_2 + 2A)^{-1}]^{-1} = \begin{bmatrix} -1 & 2 \\ 4 & 5 \end{bmatrix}^{-1}$. Therefore we have successively

$$I_2 + 2A = \frac{1}{(-1)\cdot 5 - 2\cdot 4}\begin{bmatrix} 5 & -2 \\ -4 & -1 \end{bmatrix} \Leftrightarrow I_2 + 2A = -\frac{1}{13}\begin{bmatrix} 5 & -2 \\ -4 & -1 \end{bmatrix} \Leftrightarrow$$

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

$$\Leftrightarrow 2A = \begin{bmatrix} -\frac{18}{13} & \frac{2}{13} \\ \frac{4}{13} & -\frac{12}{13} \end{bmatrix} \Leftrightarrow A = \frac{1}{2}\begin{bmatrix} -\frac{18}{13} & \frac{2}{13} \\ \frac{4}{13} & -\frac{12}{13} \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow A = \begin{bmatrix} -\frac{1}{2}\cdot\frac{18}{13} & \frac{1}{2}\cdot\frac{2}{13} \\ \frac{1}{2}\cdot\frac{4}{13} & -\frac{1}{2}\cdot\frac{12}{13} \end{bmatrix} \Leftrightarrow A = \begin{bmatrix} -\frac{9}{13} & \frac{1}{13} \\ \frac{2}{13} & -\frac{6}{13} \end{bmatrix}.$$

2. If $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, show that $A^2 = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix}$ and $A^3 = \begin{bmatrix} \cos 3\theta & -\sin 3\theta \\ \sin 3\theta & \cos 3\theta \end{bmatrix}$.

Solution: Indeed,

$$A^2 = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos^2\theta - \sin^2\theta & -2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & \cos^2\theta - \sin^2\theta \end{bmatrix} = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix}.$$

$$\begin{aligned}
A^3 = AA^2 &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta\cos 2\theta - \sin\theta\sin 2\theta & -(\cos\theta\sin 2\theta + \sin\theta\cos 2\theta) \\ \sin\theta\cos 2\theta + \cos\theta\sin 2\theta & -\sin\theta\sin 2\theta + \cos\theta\cos 2\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos(\theta+2\theta) & -\sin(\theta+2\theta) \\ \sin(\theta+2\theta) & \cos(\theta+2\theta) \end{bmatrix} \\
&= \begin{bmatrix} \cos 3\theta & -\sin 3\theta \\ \sin 3\theta & \cos 3\theta \end{bmatrix}.
\end{aligned}$$

1.5 Elementary matrices and a method for finding A^{-1}

Definition 1.5.1 An $n \times n$ matrix is called an *elementary matrix* if it can be obtain from the $n \times n$ identity matrix I_n by performing a single elementary row operation.

Example 1.5.2

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{3r_1 + r_3 \to r_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} = E - \text{elementary matrix.}$$

If $A = \begin{bmatrix} -1 & 2 & 1 & 3 \\ 2 & 1 & 6 & -3 \\ 4 & 2 & -1 & 0 \end{bmatrix}$, then $EA = \begin{bmatrix} -1 & 2 & 1 & 3 \\ 2 & 1 & 6 & -3 \\ 1 & 8 & 2 & 9 \end{bmatrix} \xleftarrow{r_3 \leftarrow r_3 + r_1} \begin{bmatrix} -1 & 2 & 1 & 3 \\ 2 & 1 & 6 & -3 \\ 4 & 2 & -1 & 0 \end{bmatrix} = A.$

Theorem 1.5.3 *If the elementary matrix E results from performing a certain row operation on I_m and A is an $m \times n$ matrix, then the product EA is the matrix that results when the same row operation is performed on A.*

If an elementary operation is performed on the identity matrix I to obtain an elementary matrix E, then there is another operation, called *the corresponding inverse operation*, when apply to E to obtain I back again.

Direct Operation	**Corresponding Inverse operation**
Multiply row i by $c \neq 0$ (cr_i)	Multiply row i by $\frac{1}{c}$ $(\frac{1}{c}r_i)$
Interchange rows i and j $(i \leftrightarrow j)$	Interchange rows i and j $(i \leftrightarrow j)$
Add c times row i to row j $(cr_i + r_j \to r_j)$	Add $-c$ times row i to row j $(-cr_i + r_j \to r_j)$

Theorem 1.5.4 *Every elementary matrix is invertible, and the inverse is also an elementary matrix.*

Proof. If $I \overset{dir.\,op.}{\longrightarrow} E \overset{inv.\,op.}{\longrightarrow} I$. Dacă $I \overset{inv.\,op.}{\longrightarrow} E_0$, then $E_0E = I$ and $EE_0 = I$, namely E is invertible.□

Theorem 1.5.5 *If A is an $n \times n$ matrix, then the following statements are equivalent:*

1. *A is invertible;*
2. *The homogeneous linear system $AX = O$ has only the trivial solution;*
3. *The reduced row-echelon form of A is I_n;*
4. *A is expressible as a product of elementary matrices.*

The last statement, (4), of theorem 1.6.7 is of particular importance since it provide us a method of finding the reduced row-echelon form of a matrix and a method of finding the inverse of an invertible matrix.

Let A be a matrix and R be its reduced row-echelon form, namely

$$A \overset{op_1}{\to} A_1 \overset{op_2}{\to} A_2 \overset{op_3}{\to} \cdots \overset{op_n}{\to} R,$$

then $R = E_n \cdot \ldots \cdot E_1 A$, where $I \overset{op_1}{\to} E_1, I \overset{op_2}{\to} E_2, \cdots, I \overset{op_n}{\to} E_n$. Therefore

$$A = (E_n \cdot \ldots \cdot E_1)^{-1} R = E_1^{-1} \cdot \ldots \cdot E_n^{-1} R.$$

If A is particularly invertible, then $R = I$, such that $A = E_1^{-1} \cdot \ldots \cdot E_n^{-1}$ and implicitly

$$A^{-1} = E_n \cdot \ldots \cdot E_1.$$

Thus, in order to find the inverse of an invertible matrix A, we first observe that we have the general property $A[B \vdots C] = [AB \vdots AC]$ and that we have successively:

$$[A \vdots I] \overset{op_1}{\to} E_1[A \vdots I] = [E_1A \vdots E_1I] \overset{op_2}{\to} E_2[E_1A \vdots E_1I] = [E_2E_1A \vdots E_2E_1I] \overset{op_3}{\to} \cdots \overset{op_n}{\to} [\underbrace{E_n \cdots E_1 A}_{=I} \vdots \underbrace{E_n \cdots E_1 I}_{=A^{-1}}].$$

Examples 1.5.6 1. Find the inverse of the matrix $A = \begin{bmatrix} -2 & -1 & 4 \\ 3 & 1 & -7 \\ 2 & 0 & -5 \end{bmatrix}$ and write A^{-1} as a product of elementary matrices;

2. Express the matrix

$$A = \begin{bmatrix} 0 & 1 & 7 & 8 \\ 1 & 3 & 3 & 8 \\ -2 & -5 & 1 & -8 \end{bmatrix}$$

in the form $A = EFGR$, where E, F, G are elementary matrices and R in row-echelon form.

(1)

$$\left[\begin{array}{ccc|ccc} -2 & -1 & 4 & 1 & 0 & 0 \\ 3 & 1 & -7 & 0 & 1 & 0 \\ 2 & 0 & -5 & 0 & 0 & 1 \end{array}\right] \xrightarrow{r_2 + r_1 \to r_1} \left[\begin{array}{ccc|ccc} 1 & 0 & -3 & 1 & 1 & 0 \\ 3 & 1 & -7 & 0 & 1 & 0 \\ 2 & 0 & -5 & 0 & 0 & 1 \end{array}\right] \xrightarrow[-2r_1 + r_3 \to r_3]{-3r_1 + r_2 \to r_2}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -3 & 1 & 1 & 0 \\ 0 & 1 & 2 & -3 & -2 & 0 \\ 0 & 0 & 1 & -2 & -2 & 1 \end{array}\right] \xrightarrow[3r_3 + r_1 \to r_1]{-2r_3 + r_2 \to r_2} \Bigg[\begin{array}{ccc|} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \underbrace{\begin{array}{ccc} -5 & -5 & 3 \\ 1 & 2 & -2 \\ -2 & -2 & 1 \end{array}}_{=A^{-1}}\Bigg].$$

Therefore $E_5 E_4 E_3 E_2 E_1 A = I_3$, where

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{r_2 + r_1 \to r_1} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} =: E_1$$

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{-3r_1 + r_2 \to r_2} \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} =: E_2$$

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{-2r_1 + r_3 \to r_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} =: E_3$$

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{-2r_3 + r_2 \to r_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} =: E_4$$

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{3r_3 + r_1 \to r_1} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} =: E_5.$$

(2)

$$A = \begin{bmatrix} 0 & 1 & 7 & 8 \\ 1 & 3 & 3 & 8 \\ -2 & -5 & 1 & -8 \end{bmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{bmatrix} 1 & 3 & 3 & 8 \\ 0 & 1 & 7 & 8 \\ -2 & -5 & 1 & -8 \end{bmatrix} \xrightarrow{2r_1 + r_3 \to r_3}$$

$$\xrightarrow{2r_1 + r_3 \to r_3} \begin{bmatrix} 1 & 3 & 3 & 8 \\ 0 & 1 & 7 & 8 \\ 0 & 1 & 7 & 8 \end{bmatrix} \xrightarrow{-r_2 + r_3 \to r_3} \begin{bmatrix} 1 & 3 & 3 & 8 \\ 0 & 1 & 7 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix} = R\text{−matrix in row-echelon form.}$$

Consequently $R = E_3 E_2 E_1 A$, or equivalently $A = E_1^{-1} E_2^{-1} E_3^{-1} R$, where

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} =: E_1 = E_1^{-1}$$

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{array}{l} 2r_1 + r_3 \to r_3 \nearrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} =: E_2 \\ -2r_1 + r_3 \to r_3 \searrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} =: E_2^{-1} \end{array}$$

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{array}{l} -r_2 + r_3 \to r_3 \nearrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} =: E_3 \\ r_2 + r_3 \to r_3 \searrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} =: E_3^{-1} \end{array} .$$

Now it is enough to take $E = E_1^{-1}, F = E_2^{-1}, G = E_3^{-1}$.

1.6 Further results on systems of equations and invertibility

Theorem 1.6.1 *Every system of linear equations has either no solution, exactly one solution or infinitely many solutions.*

Theorem 1.6.2 *If A is an invertible $n \times n$ matrix, then for each $n \times 1$ matrix b, the system of linear equations $AX = B$ has exactly one solution, namely $X = A^{-1}B$.*

Example 1.6.3 Find the solution set of the following linear system:

$$\begin{array}{rcrcrcr} -2x & - & y & + & 4z & = & 1 \\ 3x & + & y & - & 7z & = & -1 \\ 2x & & & - & 5z & = & 0 \end{array}$$

Solution: We first observe that the coefficient matrix of the given linear system is $A = \begin{bmatrix} -2 & -1 & 4 \\ 3 & 1 & -7 \\ 2 & 0 & -5 \end{bmatrix}$, which is invertible and its inverse is $A^{-1} = \begin{bmatrix} -5 & -5 & 3 \\ 1 & 2 & -2 \\ -2 & -2 & 1 \end{bmatrix}$. The given linear system can be written in the form $AX = B$, where $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and $B = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$. Consequently the unique solution of the given linear system is $\mathbf{x} = A^{-1}B = \begin{bmatrix} -5 & -5 & 3 \\ 1 & 2 & -2 \\ -2 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$.

Frequently, one have to solve a sequence of linear systems

$$AX = B_1, AX = B_2, \ldots, AX = B_k$$

each of which has the same square matrix A. If A is invertible then the solutions are

$$X_1 = A^{-1}B_1, X_2 = A^{-1}B_2, \ldots, X_k = A^{-1}B_k.$$

Otherwise the method of solving those systems, which works both for A invertible or A non-invertible, consists in forming the matrix

$$[A \vdots B_1 \vdots B_2 \vdots \cdots \vdots B_k]$$

and by reducing it to the reduced row-echelon form.

Example 1.6.4 Solve the following linear systems:

$$\begin{array}{rcrcrcr} -2x & - & y & + & 4z & = & 1 \\ 3x & + & y & - & 7z & = & -1 \\ 2x & & & - & 5z & = & 0 \end{array}, \quad \begin{array}{rcrcrcr} -2x & - & y & + & 4z & = & 2 \\ 3x & + & y & - & 7z & = & -2 \\ 2x & & & - & 5z & = & 3 \end{array}$$

Theorem 1.6.5 *Let A be a square matrix.*

1. *If B is a square matrix satisfying $BA = I$, then $B = A^{-1}$;*

2. *If B is a square matrix satisfying $AB = I$, then $B = A^{-1}$.*

Theorem 1.6.6 ***Theorem 1.6.7*** *If A is an $n \times n$ matrix, then the following statements are equivalent:*

1. *A is invertible (*);*

2. *The homogeneous linear system $AX = O$ has only the trivial solution;*

3. *The reduced row-echelon form of A is I_n;*

4. *A is expressible as a product of elementary matrices.*

5. *$AX = B$ is consistent for every $n \times 1$ matrix B;(*)*

6. *$AX = B$ has exactly one solution for every $n \times 1$ matrix B(*).*

Theorem 1.6.8 *Let A and B be square matrices of the same size. If AB is invertible, then A and B are both invertible.*

Fundamental Problem 1.6.9 Let A be a fixed $m \times n$ matrix. Find all $m \times 1$ matrices B such that the system of equations $AX = B$ is consistent.

Example 1.6.10 What conditions must a, b_1, b_2, b_3 satisfy in order for the system of the linear system

$$\begin{array}{rcrcrl} x & + & y & - & 2z & = b_1 \\ x & - & 2y & + & z & = b_2 \\ -2x & + & y & + & z & = b_3 \end{array}$$

to be consistent ?

Solution: For the augmented matrix of the given linear system we have successively:

$$\left[\begin{array}{rrrcc} 1 & 1 & -2 & \vdots & b_1 \\ 1 & -2 & 1 & \vdots & b_2 \\ -2 & 1 & 1 & \vdots & b_3 \end{array}\right] \xrightarrow[2r_1 + r_3 \to r_3]{-r_1 + r_2 \to r_2} \left[\begin{array}{rrrcc} 1 & 1 & a & \vdots & b_1 \\ 0 & -3 & 3 & \vdots & b_2 - b_1 \\ 0 & 3 & -8 & \vdots & b_3 + 2b_1 \end{array}\right] \xrightarrow{r_1 + r_2 \to r_2}$$

$$\xrightarrow{r_1 + r_2 \to r_2} \left[\begin{array}{rrrcc} 1 & 1 & a & \vdots & b_1 \\ 0 & -3 & 3 & \vdots & b_2 - b_1 \\ 0 & 0 & 0 & \vdots & b_3 + b_2 + b_1 \end{array}\right].$$

Thus, the given linear system is consistent if and only if $b_3 + b_2 + b_1 = 0$.$\square$

1.7 Diagonal, Triangular and Symmetric Matrices

Definition 1.7.1 A square matrix in which all the entries off the diagonal are zero is called a *diagonal matrix.* The general form a of a diagonal matrix is

$$D = \begin{bmatrix} d_1 & & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}.$$

Remarks 1.7.2 1. The diagonal matrix

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

is invertible iff all of its diagonal entries are non-zero. In this case

$$D^{-1} = \begin{bmatrix} \frac{1}{d_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{d_n} \end{bmatrix}.$$

2. Powers of diagonal matrices are easy to be computed. More precisely, if

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \text{ then } D^k = \begin{bmatrix} d_1^k & 0 & \cdots & 0 \\ 0 & d_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^k \end{bmatrix}.$$

3. If $A = [a_{ij}]$ is an $m \times n$ matrix, then

$$\begin{bmatrix} \delta_1 & 0 & \cdots & 0 \\ 0 & \delta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \delta_m \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \delta_1 a_{11} & \delta_1 a_{12} & \dots & \delta_1 a_{1n} \\ \delta_2 a_{21} & \delta_2 a_{22} & \dots & \delta_2 a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_m a_{m1} & \delta_m a_{m2} & \dots & \delta_m a_{mn} \end{bmatrix}$$

and

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} d_1 & & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} = \begin{bmatrix} a_{11} d_1 & a_{12} d_2 & \dots & a_{1n} d_n \\ a_{21} d_1 & a_{22} d_2 & \dots & a_{2n} d_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} d_1 & a_{m2} d_2 & \dots & a_{mn} d_n \end{bmatrix}.$$

Examples 1.7.3 Find a diagonal matrix A satisfying

$$A^{-2} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Solution: We have successively:

$$A^{-2} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Leftrightarrow [A^{-2}]^{-1} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \Leftrightarrow$$

$$\Leftrightarrow A^2 = \begin{bmatrix} \frac{1}{9} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 \end{bmatrix} \Leftrightarrow A^2 = \begin{bmatrix} \left(\frac{1}{3}\right)^2 & 0 & 0 \\ 0 & \left(\frac{1}{2}\right)^2 & 0 \\ 0 & 0 & 1^2 \end{bmatrix}.$$

Thus, such a matrix is

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

Definition 1.7.4 A square matrix in which all the entries above the main diagonal are zero is called *lower triangular* and a square matrix in which all the entries below the main diagonal are zero is called *upper triangular*. A matrix which is either lower triangular or upper triangular is called *triangular*. The general form a of a lower triangular matrix is

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

and the general form a of a upper triangular matrix is

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}.$$

Remark 1.7.5

1. A square matrix $A = [a_{ij}]$ is upper triangular iff $a_{ij} = 0$ for $i > j$;
2. A square matrix $A = [a_{ij}]$ is lower triangular iff $a_{ij} = 0$ for $i < j$.

Theorem 1.7.6

1. *The transpose of a lower triangular matrix is upper triangular, and the transpose of an upper triangular matrix is lower triangular;*
2. *The product of lower triangular matrices is a lower triangular matrix;*

3. *The product of upper triangular matrices is a upper triangular matrix;*

4. *A triangular matrix is invertible if and only if its diagonal entries are all nonzero;*

5. *The inverse of a lower triangular matrix is lower triangular and the inverse of an upper triangular matrix is upper triangular.*

Example 1.7.7 If

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 3 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 5 \end{bmatrix},$$

show that AB is also upper triangular and find $(AB)^{-1}$ and $B^{-1}A^{-1}$.

Solution:

$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 3 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 9 \\ 0 & -2 & 19 \\ 0 & 0 & 5 \end{bmatrix}.$$

$$\left[\begin{array}{ccc|ccc} 2 & 4 & 9 & 1 & 0 & 0 \\ 0 & -2 & 19 & 0 & 1 & 0 \\ 0 & 0 & 5 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\frac{1}{2}R_1, -\frac{1}{2}R_2, \frac{1}{5}R_3} \left[\begin{array}{ccc|ccc} 1 & 2 & \frac{9}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & -\frac{19}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{5} \end{array}\right] \xrightarrow[\frac{19}{2}R_3 + R_2 \to R_2]{-2R_2 + R_1 \to R_1}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & \frac{9}{2} & \frac{1}{2} & 1 & 0 \\ 0 & 1 & 0 & 0 & -\frac{1}{2} & \frac{19}{10} \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{5} \end{array}\right] \xrightarrow{-\frac{9}{2}R_3 + R_1 \to R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 1 & -\frac{9}{10} \\ 0 & 1 & 0 & 0 & -\frac{1}{2} & \frac{19}{10} \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{5} \end{array}\right].$$

Thus

$$(AB)^{-1} = B^{-1}A^{-1} = \begin{bmatrix} \frac{1}{2} & 1 & -\frac{9}{10} \\ 0 & -\frac{1}{2} & \frac{19}{10} \\ 0 & 0 & \frac{1}{5} \end{bmatrix}.$$

Definition 1.7.8 A square matrix A is called *symmetric* if $A^T = A$, or equivalently $(A)_{ij} = (A)_{ji}$

Theorem 1.7.9 *If A, B are symmetric matrices of the same size and k is any scalar, then:*

1. *A^T is symmetric;*

2. *$A + B$ and $A - B$ are symmetric;*

3. *kA is symmetric.*

Theorem 1.7.10 1. *The product of two symmetric matrices is symmetric if and only if the matrices commute;*

2. If A is an invertible symmetric matrix, then A^{-1} is symmetric.

Theorem 1.7.11 1. If A is an arbitrary matrix, then AA^T and A^TA are symmetric matrices;

2. If A is an invertible matrix, then AA^T and A^TA are invertible matrices.

Example 1.7.12 Find the values of $a, b, c \in \mathbf{R}$ for which the matrix A is symmetric, where

$$A = \begin{bmatrix} 2 & a+b+c & 4b-3c \\ 5 & 1 & 2a+c \\ 2 & 4 & 6 \end{bmatrix}.$$

Solution: The required values are solutions of the linear system $\begin{matrix} a+b+c=5 \\ 4b-3c=2 \\ 2a+c=4 \end{matrix}$ that can be solved by performing row-reduction operations in the corresponding augumented matrix, namely: $\begin{bmatrix} 1 & 1 & 1 & \vdots & 5 \\ 0 & 4 & -3 & \vdots & 2 \\ 2 & 0 & 1 & \vdots & 4 \end{bmatrix} \xrightarrow{-2r_1+r_3 \to r_3} \begin{bmatrix} 1 & 1 & 1 & \vdots & 5 \\ 0 & 4 & -3 & \vdots & 2 \\ 0 & -2 & -1 & \vdots & -6 \end{bmatrix} \xrightarrow{\frac{1}{4}r_2} \xrightarrow{\frac{1}{4}r_2}$
$\begin{bmatrix} 1 & 1 & 1 & \vdots & 5 \\ 0 & 1 & -\frac{3}{4} & \vdots & \frac{1}{2} \\ 0 & -2 & -1 & \vdots & -6 \end{bmatrix} \xrightarrow{2r_2+r_3 \to r_3} \begin{bmatrix} 1 & 1 & 1 & \vdots & 5 \\ 0 & 1 & -\frac{3}{4} & \vdots & \frac{1}{2} \\ 0 & 0 & -\frac{5}{2} & \vdots & -5 \end{bmatrix} \xrightarrow{-\frac{2}{5}r_3} \begin{bmatrix} 1 & 0 & \frac{7}{4} & \vdots & \frac{9}{2} \\ 0 & 1 & -\frac{3}{4} & \vdots & \frac{1}{2} \\ 0 & 0 & 1 & \vdots & 2 \end{bmatrix}$. The corresponding linear system of the last matrix in the above row-reduction process, which is equivalent with the initial one, has the unique solution $a = 1, b = c = 2$.

Chapter 2

Determinants

2.1 The determinant function

Definition 2.1.1 A *permutation* of the set $\{1, 2, \ldots, n\}$ is a bijective function $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$, namely an arrangement of the integers $1, 2, \ldots, n$ without omissions or repetitions. It is also written

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

An *inversion* of the permutation σ is a pair (i, j) such that $i < j$ and $\sigma(i) > \sigma(j)$. Denote by $m(\sigma)$ the number of inversions of σ and by $\varepsilon(\sigma)$ the *signature* $(-1)^{\varepsilon(\sigma)}$ of σ. A permutation σ is called *even/odd* if $m(\sigma)$ is even/odd, or equivalently $\varepsilon(\sigma) = +1/\varepsilon(\sigma) = -1$.

Definition 2.1.2 The *determinant* of a square $n \times n$ matrix $A = [a_{ij}]$ is defined to be

$$\det(A) = \sum \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdot \ldots \cdot a_{n\sigma(n)}.$$

It is denoted either by $\det(A)$ or by

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}.$$

Remarks 2.1.3 1. If $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is a 2×2 square matrix, then $\det(A) = a_{11}a_{22} - a_{12}a_{21}$.

Indeed, the only 2×2 permutations are $e = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ and $\sigma = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ and $\varepsilon(e) = 1$,

$\varepsilon(\sigma) = -1$ since $m(e) = 0$ and $m(\sigma) = 1$. Thus

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \varepsilon(e)a_{1e(1)}a_{2e(2)} + \varepsilon(\sigma)a_{1\sigma(1)}a_{2\sigma(2)} = a_{11}a_{22} - a_{12}a_{21}.$$

For example $\begin{vmatrix} 1 & 2 \\ -1 & 3 \end{vmatrix} = 1 \cdot 3 - (-1)3 = 3 + 3 = 6.$

2. A 2×2 matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is invertible if and only if $\det(A) \neq 0$. In this case

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

3. If $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ is a 2×2 square matrix, then

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} =$$

$$= \begin{array}{|ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

$$- \quad - \quad - \quad + \quad + \quad +$$

Indeed, the only 3×3 square matrices are

$$e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \sigma_1 := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$

$$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

and $\varepsilon(e) = \varepsilon(\sigma_4) = \varepsilon(\sigma_5) = 1$ while $\varepsilon(\sigma_1) = \varepsilon(\sigma_2) = \varepsilon(\sigma_4) = -1$ since $m(e) = 0, m(\sigma_4) = m(\sigma_5) = 2$ and $m(\sigma_1) = 1, m(\sigma_2) = 3, m(\sigma_3) = 1$.

For example $\begin{vmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ -2 & 2 & 3 \end{vmatrix} = \begin{array}{|ccc|cc} 1 & -1 & 2 & 1 & -1 \\ 0 & 2 & 1 & 0 & 2 \\ -2 & 2 & 3 & -2 & 2 \end{array} = 6 + 2 + 0 - (-8) - 2 - 0 = 8 + 8 = 16.$

2.2 Evaluating determinants by row reduction

Theorem 2.2.1 *Let A be a square matrix.*

1. *If A has a row of zeros, then* $\det(A) = 0$.

2. $\det(A) = \det(A^T)$.

Theorem 2.2.2 *If A is an $n \times n$ triangular matrix, then* $\det(A)$ *is the product of the entries on the main diagonal, namely* $\det(A) = a_{11}a_{22} \cdot \ldots \cdot a_{nn}$

Theorem 2.2.3 *Let A be an $n \times n$ square matrix.*

1. *If B is the matrix that results when a single row or a single column of A is multiplied by a scalar k, then* $\det(B) = k\det(A)$*; Consequently* $\det(kA) = k^n \det(A)$.

2. *If B is the matrix that results when two rows or two columns of A are interchanged, then* $\det(B) = -\det(A)$*;*

3. *If B is the matrix that results when a multiple of one row of A is added to another row or when a multiple of one column of A is added to another column, then* $\det(B) = \det(A)$.

Corollary 2.2.4 *If A is a square matrix with two proportional rows or two proportional columns, then* $(A) = 0$.

Example 2.2.5 Find the determinant of $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$.

Solution:

$$\det(A) = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{vmatrix} \underset{-r_1 + r_2 \to r_2}{=}$$

$$= \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -2 & -2 & -4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{vmatrix} = (-2)\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{vmatrix} = (-2)\begin{vmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{vmatrix} =$$

$$= -2(2 - 1 + 1) = -4.$$

2.3 Properties of determinant function

Theorem 2.3.1 *Let A, B and C be $n\times n$ matrices that differ only in a single row, say the r^{th}, and assume that the r^{th} row of C can be obtained by adding corresponding entries in the r^{th} row of A and B. Then* $\det(C) = \det(A) + \det(B)$.

In other words

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{11}+a'_{11} & a_{12}+a'_{12} & \cdots & a_{1n}+a'_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{11} & a'_{12} & \cdots & a'_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix}$$

Theorem 2.3.2 *If A and B are square matrices of the same size, then* $\det(AB) = \det(A)\det(B)$.

Theorem 2.3.3 *If A is an invertible matrix, then* $\det(A^{-1}) = \frac{1}{\det(A)}$.

Theorem 2.3.4 *If A is an $n\times n$ matrix, then the following statements are equivalent:*

1. *A is invertible;(*)*
2. *The homogeneous linear system $AX = O$ has only the trivial solution;*
3. *The reduced row-echelon form of A is I_n;*
4. *A is expressible as a product of elementary matrices.*
5. *$AX = B$ is consistent for every $n\times 1$ matrix B;*
6. *$AX = B$ has exactly one solution for every $n\times 1$ matrix B.*
7. $\det(A) \neq 0$*(*)*

Example 2.3.5 1. Consider the matrix $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ x & y & z & w \\ w & x & y & z \\ y+z & z+w & w+x & x+y \end{bmatrix}$. Show that $\det(A) = 0$ for all $x, y, z, w \in \mathbf{R}$.

Solution: For the determinant of A we have successively:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ x & y & z & w \\ w & x & y & z \\ y+z & z+w & w+x & x+y \end{vmatrix} \overset{r_2+r_4\to r_4}{=} \begin{vmatrix} 1 & 1 & 1 & 1 \\ x & y & z & w \\ w & x & y & z \\ x+y+z & y+z+w & z+w+x & w+x+y \end{vmatrix}$$

$$\overset{r_3+r_4\to r_4}{=} \begin{vmatrix} 1 & 1 & 1 & 1 \\ x & y & z & w \\ w & x & y & z \\ x+y+z+w & x+y+z+w & y+z+w+x & z+w+x+y \end{vmatrix} =$$

$$= (x+y+z+w) \begin{vmatrix} 1 & 1 & 1 & 1 \\ x & y & z & w \\ w & x & y & z \\ 1 & 1 & 1 & 1 \end{vmatrix} = 0$$

for all $x, y, z, w \in \mathbf{R}$.

2. Let

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}.$$

Assuming that $\det(A) = -7$, find $\det(3A), \det(A^{-1}), \det(2A^{-1})$ and $\det \begin{bmatrix} a & g & d \\ b & h & e \\ c & i & f \end{bmatrix}$.

Solution: • $\det(3A) = 3^3 \det(A) = 27(-7)$.

• $\det(A^{-1}) = \frac{1}{\det(A)} = \frac{1}{-7} = -\frac{1}{7}$.

• $\det(2A^{-1}) = 2^3 \det(A^{-1}) = 8\Big(-\frac{1}{7}\Big) = -\frac{8}{7}$.

• $\det \begin{bmatrix} a & g & d \\ b & h & e \\ c & i & f \end{bmatrix} = \det \begin{bmatrix} a & g & d \\ b & h & e \\ c & i & f \end{bmatrix}^T = \begin{vmatrix} a & b & c \\ g & h & i \\ d & e & f \end{vmatrix} = - \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = -(-7) = 7.$

3. Show that $\begin{vmatrix} a_1 + b_1 t & a_2 + b_2 t & a_3 + b_3 t \\ a_1 t + b_1 & a_2 t + b_2 & a_3 t + b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = (1-t^2) \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$

Solution: Indeed, we have successively:

$$\begin{vmatrix} a_1+b_1t & a_2+b_2t & a_3+b_3t \\ a_1t+b_1 & a_2t+b_2 & a_3t+b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ a_1t+b_1 & a_2t+b_2 & a_3t+b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} + \begin{vmatrix} b_1t & b_2t & b_3t \\ a_1t+b_1 & a_2t+b_2 & a_3t+b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$= \underbrace{\begin{vmatrix} a_1 & a_2 & a_3 \\ a_1t & a_2t & a_3t \\ c_1 & c_2 & c_3 \end{vmatrix}}_{=0} + \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} + \begin{vmatrix} b_1t & b_2t & b_3t \\ a_1t & a_2t & a_3t \\ c_1 & c_2 & c_3 \end{vmatrix} + \underbrace{\begin{vmatrix} b_1t & b_2t & b_3t \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}}_{=0}$$

$$= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} + t^2 \begin{vmatrix} b_1 & b_2 & b_3 \\ a_1 & a_2 & a_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} - t^2 \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$= (1-t^2) \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

2.4 Cofactor expansion. Cramer's rule

Definition 2.4.1 If A is a square matrix, then the *minor of entry* a_{ij} is denoted by M_{ij} and is defined to be the determinant of the submatrix that remains after the i^{th} row and the j^{th} column are deleted from A. The number $(-1)^{i+j}M_{ij}$ is denoted by C_{ij} and is called the *cofactor of entry* a_{ij}.

Example 2.4.2

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{12}a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ &= (-1)^{1+1}a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{1+2}a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + (-1)^{1+3}a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \end{aligned}$$

Similarly

$$
\begin{aligned}
\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \\
&= -a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{22}(a_{11}a_{33} - a_{13}a_{31}) - a_{32}(a_{11}a_{23} - a_{13}a_{21}) \\
&= (-1)^{1+2}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + (-1)^{2+2}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + (-1)^{3+2}\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \\
&= a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32}
\end{aligned}
$$

Theorem 2.4.3 If $A = [a_{ij}]$ is a square $n \times n$ matrix, then

$$
det(A) = \underbrace{a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}}_{\text{The cofactor expansion of } \det(A) \text{ along the } i^{th} \text{ row}} = \underbrace{a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}}_{\text{The cofactor expansion of } \det(A) \text{ along the } j^{th} \text{ column}} .
$$

Example 2.4.4 Show that:

$$
\begin{vmatrix} 1 & 1 & 1 & 1 \\ a_1 & a_2 & a_3 & a_4 \\ a_1^2 & a_2^2 & a_3^2 & a_4^2 \\ a_1^3 & a_2^3 & a_3^3 & a_4^3 \end{vmatrix} = (a_1 - a_2)(a_1 - a_3)(a_1 - a_4)(a_2 - a_3)(a_2 - a_4)(a_3 - a_4)
$$

If $A = [a_{ij}]$ is a square $n \times n$ matrix and C_{ij} is the cofactor of the entry a_{ij}, then the matrix

$$
\begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}
$$

is called the matrix of cofactors from A. *The transpose of this matrix is called* the adjoint of A *and is denoted by $adj(A)$.* If A is an invertible matrix, then $A^{-1} = \frac{1}{\det(A)} adj(A)$. The inverse of an invertible lower triangular matrix is lower triangular and the inverse of an invertible upper triangular matrix is upper triangular. *Find necessary and sufficient conditions on $a, b, c \in \mathbb{R}$*

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

such that the matrix A is invertible. In those cases find A^{-1}.

According to theorem 2.3.4, A is invertible if and only if $\det(A) \neq 0$. But since

$$\det(A) = V(a,b,c) = \begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = (c-a)(c-b)(b-a),$$

it follows that A is invertible if and only if $a \neq b, b \neq c$ and $c \neq a$. In this case

$$A^{-1} = \frac{1}{\det(A)} adj(A) = \frac{1}{\det(A)} \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix}^T =$$

$$= \frac{1}{(c-a)(c-b)(b-a)} \begin{bmatrix} \begin{vmatrix} b & c \\ b^2 & c^2 \end{vmatrix} & -\begin{vmatrix} a & c \\ a^2 & c^2 \end{vmatrix} & \begin{vmatrix} a & b \\ a^2 & b^2 \end{vmatrix} \\ -\begin{vmatrix} 1 & 1 \\ b^2 & c^2 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ a^2 & c^2 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ a^2 & b^2 \end{vmatrix} \\ \begin{vmatrix} 1 & 1 \\ b & c \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ a & c \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ a & b \end{vmatrix} \end{bmatrix}^T =$$

$$= \frac{1}{(c-a)(c-b)(b-a)} \begin{bmatrix} bc(c-b) & -ac(c-a) & ab(b-a) \\ -(c^2-b^2) & c^2-a^2 & -(b^2-a^2) \\ c-b & -(c-a) & b-a \end{bmatrix}^T =$$

$$= \frac{1}{(c-a)(c-b)(b-a)} \begin{bmatrix} bc(c-b) & b^2-c^2 & c-b \\ ac(a-c) & c^2-a^2 & a-c \\ ab(b-a) & a^2-b^2 & b-a \end{bmatrix} = \begin{bmatrix} \frac{bc(c-b)}{(c-a)(c-b)(b-a)} & \frac{(b-c)(b+c)}{(c-a)(c-b)(b-a)} & \frac{c-b}{(c-a)(c-b)(b-a)} \\ \frac{ac(a-c)}{(c-a)(c-b)(b-a)} & \frac{(c-a)(c+a)}{(c-a)(c-b)(b-a)} & \frac{a-c}{(c-a)(c-b)(b-a)} \\ \frac{ab(b-a)}{(c-a)(c-b)(b-a)} & \frac{(a-b)(a+b)}{(c-a)(c-b)(b-a)} & \frac{b-a}{(c-a)(c-b)(b-a)} \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{bc}{(c-a)(b-a)} & \frac{b+c}{(a-c)(b-a)} & \frac{1}{(c-a)(b-a)} \\ \frac{ac}{(c-b)(a-b)} & \frac{c+a}{(c-b)(b-a)} & \frac{1}{(c-b)(a-b)} \\ \frac{ab}{(c-a)(c-b)} & \frac{a+b}{(c-a)(b-c)} & \frac{1}{(c-a)(c-b)} \end{bmatrix}.$$

(Cramer's Rule) If $AX = B$ is a system of n linear equations in n unknowns such that $\det(A) \neq 0$, then the system has a unique solution. This solution s

$$x_1 = \frac{\det(A_1)}{\det(A)}, x_2 = \frac{\det(A_2)}{\det(A)}, \ldots, x_n = \frac{\det(A_n)}{\det(A)},$$

where A_j is the matrix obtained from A by replacing its j^{th} row with the entries of the column matrix B. *Provide necessary and sufficient conditions on $a_1, a_2, a_3, a_4 \in \mathbf{R}$ such that the linear* *system*

$$
\begin{array}{rcrcrcrl}
x & + & y & + & z & + & w & = 1 \\
ax & + & by & + & cz & + & dw & = \alpha \\
a^2x & + & b^2y & + & c^2z & + & d^2w & = \alpha^2 \\
a^3x & + & b^3y & + & c^3z & + & d^3w & = \alpha^3
\end{array}
\tag{2.1}
$$

has exactly one solution for each $\alpha \in \mathbf{R}$*. In this case solve the system by using the Cramer's rule.*

The given linear system has a unique solution for each $\alpha \in \mathbf{R}$ if and only if its coefficient matrix

$$
A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{bmatrix}
$$

is invertible, or, equivalently $\det(A) \neq 0$. But since

$$
\det(A) = V(a, b, c, d) = (d - a)(d - b)(d - c)(c - a)(c - b)(b - a),
$$

it follows that the given linear system has unique solution for each $\alpha \in \mathbb{R}$ if and only if the scalars a, b, c, d are pairwise disjoint.

By Crammer's rule, it follows that the unique solution of the system is

$$
x = \frac{V(\alpha, b, c, d)}{V(a, b, c, d)} = \frac{(d - \alpha)(c - \alpha)(b - \alpha)}{(d - a)(c - a)(b - a)};
$$

$$
y = \frac{V(a, \alpha, c, d)}{V(a, b, c, d)} = \frac{(\alpha - a)(c - \alpha)(d - \alpha)}{(b - a)(c - b)(d - b)};
$$

$$
z = \frac{V(a, b, \alpha, d)}{V(a, b, c, d)} = \frac{(\alpha - a)(\alpha - b)(d - \alpha)}{(c - a)(c - b)(d - c)};
$$

$$
w = \frac{V(a, b, c, \alpha)}{V(a, b, c, d)} = \frac{(\alpha - a)(\alpha - b)(\alpha - c)}{(d - a)(d - b)(d - c)}.
$$

Chapter 3

Euclidean vector spaces

3.1 Euclidean n-space

Definition 3.1.1 If n is a positive integer, then an *ordered n-tuple* is a sequence of n numbers $(a_1, a_2, \ldots, a_n)$. The set of $\mathbb{R}^n$ of all n-tuples is called *n-space*.

A 2-tuple is also called *ordered pair* and a 3-tuple is called *ordered triple* and both of them have geometric interpretation. Thus an n-tuple can be viewed as either an 'generalized point' or a 'generalized vector'.

Definition 3.1.2 Two vectors $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_n)$ in $\mathbb{R}^n$ are called *equal* if $u_1 = v_1, u_2 = v_2, \ldots, u_n = v_n$. Their *sum* $u+v$ is defined by $u+v = (u_1+v_1, u_2+v_2, \ldots, u_n+v_n)$ and, for a sscalar k, the *scalar multiple* is defined by $ku = (ku_1, ku_2, \ldots, ku_n)$.

The *zerovector* in $\mathbb{R}^n$ is denoted by 0 and is defined to be the vector $0 = (0, 0, \ldots, 0)$. If $u = (u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$ is any vector, the *negative* or *additive inverse* of u is denoted by $-u$ and is defined by $-u = (-u_1, -u_2, \ldots, -u_n)$. The *difference* $u - v$ of the vectors $(u_1, u_2, \ldots, u_n), v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ is defined by $u - v := u + (-v)$. In termes of components we have $u - v = (u_1 - v_1, u_2 - v_2, \ldots, u_n - v_n)$.

Theorem 3.1.3 *If $u = (u_1, u_2, \ldots, u_n), v = (v_1, v_2, \ldots, v_n), w = (w_1, w_2, \ldots, w_n) \in \mathbb{R}^n$ and k, l are scalars, then*

1. $u + v = v + u$.
2. $u + (v + w) = (u + v) + w$.
3. $u + 0 = 0 + u = u$.
4. $u + (-u) = 0$, *namely* $u - u = 0$.

5. $k(lu) = (kl)u$.

6. $k(u+v) = ku + kv$.

7. $(k+l)u = ku + lu$.

8. $1u = u$.

Definition 3.1.4 If $u = (u_1, u_2, \dots, u_n), v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ are any vectors, then the *Euclidean Product* $u \cdot v$ is defined by $u \cdot v := u_1v_1 + u_2v_2 + \cdots + u_nv_n$.

Theorem 3.1.5 *If* $u = (u_1, u_2, \dots, u_n), v = (v_1, v_2, \dots, v_n), w = (w_1, w_2, \dots, w_n) \in \mathbb{R}^n$ *and* k, l *are scalars, then*

1. $u \cdot v = v \cdot u$.

2. $k(u \cdot v) = (ku) \cdot v = u \cdot (kv)$.

3. $(u+v) \cdot w = u \cdot w + v \cdot w$.

4. $u \cdot u \leq 0$ *and* $u \cdot u = 0$ *if and only if* $u - 0$.

Definition 3.1.6 The *Euclidean norm* or the *Euclidean length* $||u||$ of a vector $u = (u_1, u_2, \dots, u_n) \in \mathbb{R}^n$ is defined by $||u|| := \sqrt{u \cdot u} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$ and the distance $d(u, v)$ between the points $u = (u_1, u_2, \dots, u_n), v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ is defined by

$$d(u,v) = ||u - v|| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}.$$

Theorem 3.1.7 *(Cauchy-Schwarz Inequality in* $\mathbb{R}^n$*) If* $u = (u_1, u_2, \dots, u_n), v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$, *then* $|u \cdot v| \leq ||u|| \cdot ||v||$. *In terms of components, the inequality becomes*

$$|u_1v_1 + u_2v_2 + \cdots + u_nv_n| \leq \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2} \cdot \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

Theorem 3.1.8 *If* $u, v \in \mathbb{R}^n$ *and* k *is any scalar, then*

1. $||u|| \geq 0$.

2. $||u|| = 0$ *if and only if* $u = 0$.

3. $||ku|| = |k|\,||u||$.

4. $||u+v|| \leq ||u|| + ||v||$ *(Triangle inequality).*

Theorem 3.1.9 *If $u, v, w \in \mathbb{R}^n$ and k is any scalar, then*

1. $d(u,v) \geq 0$.

2. $d(u,v) = 0$ *if and only if* $u = v$.

3. $d(u,v) = d(v,u)$.

4. $d(u,v) \leq d(u,w) + d(w,v)$ *(Triangle inequality).*

Theorem 3.1.10 *If $u, v \in \mathbb{R}^n$, then $u \cdot v = \frac{1}{4}||u+v||^2 - \frac{1}{4}||u-v||^2$.*

Proof. Indeed, by adding the identities

$$\begin{aligned} ||u+v||^2 &= (u+v)\cdot(u+v) = ||u||^2 + 2u\cdot v + ||v||^2 \\ ||u-v||^2 &= (u-v)\cdot(u-v) = ||u||^2 - 2u\cdot v + ||v||^2, \end{aligned}$$

we immediately get the required identity.□

Definition 3.1.11 Two vectors $u, v \in \mathbb{R}^n$ are said to be *orthogonal* if $u \cdot v = 0$.

Theorem 3.1.12 *(Theorem of Pythagoras for $\mathbb{R}^n$) If $u, v \in \mathbb{R}^n$ are orthogonal vectors, then the equality $||u+v||^2 = ||u||^2 + ||v||^2$ holds.*

Proof. Indeed, we have successively:

$$||u+v||^2 = (u+v)\cdot(u+v) = ||u||^2 + 2u\cdot v + ||v||^2 = ||u||^2 + ||v||^2.\square$$

A vector $u = (u_1, \ldots, u_n) \in \mathbb{R}^n$ can be naturally identified with a column or with a row matrix, namely with

$$\underline{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \text{ or } \overline{u} = [u_1 \; u_2 \cdots \; u_n].$$

Consequently, the set $\mathbb{R}^n$ of n-tuples can be naturally identified with either the space of column matrices, or with the space of row matrices. The sum $u + v$ of two vectors $u = (u_1, \ldots, u_n), v = (v_1, \ldots, v_n)$ will be identified consquently identified with the sum

$$\underline{u} + \underline{v} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$$

of the column matrices $\underline{u}$ and $\underline{v}$, or with the sum

$$\overline{u}+\overline{v}=[u_1\ u_2\cdots\ u_n]+[v_1\ v_2\cdots\ v_n]=[u_1+v_1\ u_2+v_2\cdots\ u_n+v_n]$$

of the row matrices $\overline{u}$ and $\overline{v}$.

Similarly, the scalar multiple ku of the real k with the ordered n-tuple $u=(u_1,u_2,\cdots,u_n)$ can be either identified with the scalar multiple

$$k\underline{u}=k\begin{bmatrix}u_1\\u_2\\\vdots\\u_n\end{bmatrix}=\begin{bmatrix}ku_1\\ku_2\\\vdots\\ku_n\end{bmatrix},$$

or with the scalar multiple

$$k\overline{u}=k[u_1,u_2,\ldots,u_n]=[ku_1,ku_2\ldots,ku_n].$$

3.2 Linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$

Definition 3.2.1 A mapping $T_A:\mathbb{R}^n\to\mathbb{R}^m, \underline{T}_A(\underline{x})=A\underline{x}$, where A is an $m\times n$ matrix, is called *linear transformation*, or *linear operator*, if $m=n$. The linear transformation T_A is equally called the *multiplication by A. If $A=[a_{ij}]_{mn}$, then*

$$T\left(\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\right)=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n\\\vdots\\a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n\end{bmatrix}.$$

Observe that

$$T\left(\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}\right)=\begin{bmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{bmatrix},T\left(\begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix}\right)=\begin{bmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{bmatrix},\ldots,T\left(\begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}\right)=\begin{bmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{bmatrix}$$

The matrix A is called *the standard matrix* of the linear transformation T and it is usually denoted by $[T]$ and the multiplication by A mapping is also denoted by T_A. Therefore $[T]=$

$[T(e_1)\vdots T(e_2)\vdots\cdots\vdots T(e_n)]$, where

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

1. The *orthogonal projection of* $\mathbb{R}^2$ *on the x-axis*, $p_x : \mathbb{R}^2 \to \mathbb{R}^2, p_x(x_1, x_2) = (x_1, 0)$ is a linear mapping since

$$\underline{p}_x\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

and $[p_x] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Its equations are : $\begin{array}{rcl} w_1 & = & x_1 \\ w_2 & = & 0. \end{array}$

2. The *orthogonal projection of* $\mathbb{R}^2$ *on the y-axis*, $p_y : \mathbb{R}^2 \to \mathbb{R}^2, p_y(y_1, y_2) = (0, y_2)$ is a linear mapping since

$$\underline{p}_y\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

and $[p_y] = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Its equations are : $\begin{array}{rcl} w_1 & = & 0 \\ w_2 & = & y_2. \end{array}$

3. The *reflection of* $\mathbb{R}^3$ *about the* x_1x_2*-plane*, $S_{x_1x_2} : \mathbb{R}^3 \to \mathbb{R}^3, S_{x_1x_2}(x, y, z) = (x, y, -z)$ is a linear mapping since

$$\underline{S}_{x_1x_2}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ y \\ -z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

and $[S_{x_1x_2}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. Its equations are : $\begin{array}{rcl} w_1 & = & x \\ w_2 & = & y \\ w_3 & = & -z \end{array}$

4. The *reflection of* $\mathbb{R}^3$ *about the* x_1x_3*-plane*, $S_{x_1x_3} : \mathbb{R}^3 \to \mathbb{R}^3, S_{x_1x_3}(x, y, z) = (x, -y, z)$ is a linear mapping since

$$\underline{S}_{x_1x_3}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ -y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

and $[S_{x_1x_3}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Its equations are : $\begin{matrix} w_1 & = & x \\ w_2 & = & -y \\ w_3 & = & z \end{matrix}$

5. The *reflection of* $\mathbb{R}^3$ *about the* x_2x_3*-plane*, $S_{x_2x_3} : \mathbb{R}^3 \to \mathbb{R}^3, S_{x_2x_3}(x,y,z) = (x,-y,z)$ is a linear mapping since

$$\underline{S}_{x_2x_3}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} -x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

and $[S_{x_2x_3}] = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Its equations are : $\begin{matrix} w_1 & = & -x \\ w_2 & = & y \\ w_3 & = & z \end{matrix}$

6. The *orthogonal projection of* $\mathbb{R}^3$ *on the* x_1x_2*-plane*, $P_{x_1x_2} : \mathbb{R}^3 \to \mathbb{R}^3, P_{x_1x_2}(x,y,z) = (x,y,0)$ is a linear mapping since

$$\underline{P}_{x_1x_2}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

and $[P_{x_1x_2}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Its equations are : $\begin{matrix} w_1 & = & x \\ w_2 & = & y \\ w_3 & = & 0 \end{matrix}$

7. The *orthogonal projection of* $\mathbb{R}^3$ *on the* x_1x_3*-plane*, $P_{x_1x_3} : \mathbb{R}^3 \to \mathbb{R}^3, P_{x_1x_3}(x,y,z) = (x,0,z)$ is a linear mapping since

$$\underline{P}_{x_1x_3}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ 0 \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

and $[P_{x_1x_3}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Its equations are : $\begin{matrix} w_1 & = & x \\ w_2 & = & 0 \\ w_3 & = & z \end{matrix}$

8. The *orthogonal projection of* $\mathbb{R}^3$ *on the* x_2x_3*-plane*, $P_{x_2x_3} : \mathbb{R}^3 \to \mathbb{R}^3, P_{x_2x_3}(x,y,z) = (0,y,z)$ is a linear mapping since

$$\underline{P}_{x_2x_3}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 0 \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

and $[P_{x_2 x_3}] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Its equations are: $\begin{aligned} w_1 &= 0 \\ w_2 &= y \\ w_3 &= z \end{aligned}$

9. The *rotation operator of* $\mathbb{R}^2$ *through a fixed angle* θ,

$$R_\theta : \mathbb{R}^2 \to \mathbb{R}^2, R_\theta = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)$$

is a linear mapping since

$$\underline{R}_\theta\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

and $[R_\theta] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Its equations are: $\begin{aligned} w_1 &= x\cos\theta - y\sin\theta \\ w_2 &= x\sin\theta + y\cos\theta. \end{aligned}$

Indeed, the rotation operator R_θ rotates the point $(x, y) = (r\cos\varphi, r\sin\varphi)$, counterclockwise with the angle θ if $\theta > 0$ and clockwise if $\theta < 0$, the coordinates of the rotated point being $(w_1, w_2) = (r\cos(\theta + \varphi), r\sin(\theta + \varphi))$, such that one gets $\begin{aligned} w_1 &= x\cos\theta - y\sin\theta \\ w_2 &= x\sin\theta + y\cos\theta. \end{aligned}$

10. The *rotation operator of* $\mathbb{R}^3$ *through a fixed angle* θ *about an oriented axis*, rotates about the axis of rotation each point of $\mathbb{R}^3$ in such a way that its associated vector sweeps out some portion of the cone determine by the vector itself an by a vector which gives the direction and the orientation of the considered oriented axis. The angle of the rotattion is measured at the base of the cone and it is measured clockwise or counterclockwise in relation with a viewpoint along the axis looking toward the origin. As in $\mathbb{R}^2$, the positives angles generates counterclockwise roattions and negative angles generates clockwise roattions. The counterclockwise sense of rotaion can be determined by the right-hand rule: If the thumb of the right hand points the direction of the direction of the oriented axis, then the cupped fingers points in a counterclockwise direction. The rotation operators in $\mathbb{R}^3$ are linear. For example

    (a) The counterclockwise rotation about the positive x-axis through an angle θ has the equations $\begin{aligned} w_1 &= x \\ w_2 &= y\cos\theta - z\sin\theta \\ w_3 &= y\sin\theta + z\cos\theta \end{aligned}$, its standard matrix is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$.

    (b) The counterclockwise rotation about the positive y-axis through an angle θ has the equations $\begin{aligned} w_1 &= x\cos\theta + z\sin\theta \\ w_2 &= y \\ w_3 &= -x\sin\theta + z\cos\theta \end{aligned}$, its standard matrix is $\begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}$.

(c) The counterclockwise rotation about the positive z-axis through an angle θ has the equations $\begin{aligned} w_1 &= x\cos\theta - y\sin\theta \\ w_2 &= x\sin\theta + y\cos\theta \\ w_3 &= z \end{aligned}$, its standard matrix is $\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

(d) The *homotopy* of ratio $k \in \mathbb{R}$ is the linear operator $H_k : \mathbb{R}^n \to \mathbb{R}^n, H_k(x) = kx$, Its standard matrix is

$$\begin{bmatrix} k & 0 & \cdots & 0 \\ 0 & k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & k \end{bmatrix} \text{ and its equations are } \begin{aligned} w_1 &= kx_1 \\ w_2 &= kx_2 \\ &\vdots \\ w_n &= kx_n. \end{aligned}$$

If $0 \leq k \leq 1$, then H_k is called *contraction*, and for $k \geq 1, H_k$ is called *dilatation.*

• Let A, B be matrices of sizes $n \times k$ and $k \times m$ respectively, and $T_A : \mathbb{R}^n \to \mathbb{R}^k, T_B : \mathbb{R}^k \to \mathbb{R}^m$ are the linear transformations of standard matrices A and B respectively, namely

$$\underline{T}_A(\underline{x}) = A\underline{x}, \underline{T}_B(\underline{y}) = B\underline{y}.$$

Then the composed mapping $T_B \circ T_A$ is also linear and $T_B \circ T_A = T_{BA}$. Indeed

$$(T_B \circ T_A)(\underline{x}) = T_B(T_A)(\underline{x})) = T_B(A\underline{x}) = B(A\underline{x}) = (BA)\underline{x} = T_{BA}(\underline{x}).$$

We have just proved that $[T_2 \circ T_1] = [T_2][T_1]$ for any two linear mappings whose composition $T_2 \circ T_1$ is defined. Similarly, if the composed map $T_3 \circ T_2 \circ T_1$, of the linear mappings T_3,T_2,T_1, is defined, then $[T_3 \circ T_2 \circ T_1] = [T_3][T_2][T_1]$.

• Let $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ be the reflection operator about the line $y = x$ and let $T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ be the orthogonal projection about the y-axis. Find $T_1 \circ T_2, T_2 \circ T_1$ and their standard matrices.

• Show that $R_{\theta_1} \circ R_{\theta_2} = R_{\theta_2} \circ R_{\theta_1}$ and it is the rotation of angle $\theta_1 + \theta_2$.

3.3 Properties of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$

Definition 3.3.1 *A linear transformation* $T : \mathbb{R}^n \to \mathbb{R}^m$ *is said to be* one-to-one *if* $u_1, u_2 \in \mathbb{R}^n, u_1 \neq u_2 \Rightarrow Tu_1 \neq Tu_2$.

It follows that each vector w in the *range* $\{Tx : x \in \mathbb{R}^n\}$ of an one-to-one linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ there is exactly one vector x such that $Tx = w$.

Theorem 3.3.2 If A is an $n \times n$ matrix and $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is the multiplication by A, then the following statements are equivalent:

1. A is invertible.

2. The range of T_A is $\mathbb{R}^n$.

3. T_A is one-to-one.

If $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is a one-to-one linear operator, then the matrix A is invertible and $T_{A^{-1}} : \mathbb{R}^n \to \mathbb{R}^n$ is also linear. Moreover,

$$T_A(T_{A^{-1}}(\underline{\mathrm{x}})) = AA^{-1}\underline{\mathrm{x}} = \underline{\mathrm{x}} = T_I\underline{\mathrm{x}} = id_{\mathbb{R}^n}(\underline{\mathrm{x}}), \forall \underline{\mathrm{x}} \in \mathbb{R}^n.$$

$$T_{A^{-1}}(T_A(\underline{\mathrm{x}})) = A^{-1}A\underline{\mathrm{x}} = \underline{\mathrm{x}} = T_I\underline{\mathrm{x}} = id_{\mathbb{R}^n}(\underline{\mathrm{x}}), \forall \underline{\mathrm{x}} \in \mathbb{R}^n.$$

Consequently T_A is also invertible and its inverse is $T_A^{-1} = T_{A^{-1}}$. Therefore, for the standard matrix of the inverse of an one-to-one linear operator $T : \mathbb{R}^n \to \mathbb{R}^n$, we have $[T^{-1}] = [T]^{-1}$.

Examples 3.3.3 The *rotation operator of* $\mathbb{R}^2$ *through a fixed angle* θ,

$$R_\theta : \mathbb{R}^2 \to \mathbb{R}^2, R_\theta = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)$$

is a linear mapping since has the standard matric and $[R_\theta] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, which is invertible and $[R_\theta]^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} = [R_{-\theta}]$. Consequently R_θ is invertible and $R_\theta^{-1} = R_{-\theta}$.

Theorem 3.3.4 *A transformation* $T : \mathbb{R}^n \to \mathbb{R}^m$ *is linear if and only if*

1. $T(u+v) = T(u) + T(v)$ *for all* $u, v \in \mathbb{R}^n$.

2. $T(cu) = cT(u)$ *for all* $u \in \mathbb{R}^n$.

Proof. Indeed, if $T = T_A$, then $\underline{\mathrm{T}} = \underline{\mathrm{T}}_A$, namely $\underline{\mathrm{T}}_A(\underline{\mathrm{u}}+\underline{\mathrm{v}}) = A(\underline{\mathrm{u}}+\underline{\mathrm{v}}) = A\underline{\mathrm{u}} + A\underline{\mathrm{v}} = \underline{\mathrm{T}}_A\underline{\mathrm{u}} + \underline{\mathrm{T}}_A\underline{\mathrm{v}}$.

Conversely, if the given conditions are satisfied, then one can easily show that for any k vectors $u_1, \ldots, u_k \in \mathbb{R}^n$ we have $\underline{\mathrm{T}}(\underline{\mathrm{u}}_1 + \cdots + \underline{\mathrm{u}}_k) = \underline{\mathrm{T}}(\underline{\mathrm{u}}_1) + \cdots + \underline{\mathrm{T}}(\underline{\mathrm{u}}_k)$. Oncan now show that $T = T_A$, where $A = [\underline{\mathrm{T}}(\underline{\mathrm{e}}_1) \vdots \cdots \vdots \underline{\mathrm{T}}(\underline{\mathrm{e}}_n)]$, where

$$\underline{\mathrm{e}}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \underline{\mathrm{e}}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \underline{\mathrm{e}}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Indeed, for

$$\underline{\mathrm{x}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + x_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = x_1\underline{\mathrm{e}}_1 + x_2\underline{\mathrm{e}}_2 + \cdots + x_n\underline{\mathrm{e}}_n$$

we have

$$\underline{\mathrm{T}}(\underline{\mathrm{x}}) = \underline{\mathrm{T}}(x_1\underline{\mathrm{e}}_1 + x_2\underline{\mathrm{e}}_2 + \cdots + x_n\underline{\mathrm{e}}_n) = x_1\underline{\mathrm{T}}(\underline{\mathrm{e}}_1) + x_2\underline{\mathrm{T}}(\underline{\mathrm{e}}_2) + \cdots + x_n\underline{\mathrm{T}}(\underline{\mathrm{e}}_n) = A\underline{\mathrm{x}}.\square$$