INTERNATIONAL COLLEGE OF ECONOMICS AND FINANCE

# LECTURE NOTES
# IN
# LINEAR ALGEBRA

$$\begin{pmatrix} \alpha_{11} & \beta_{12} & \gamma_{13} & \delta_{14} & \varepsilon_{15} \\ \zeta_{21} & \eta_{22} & \theta_{23} & \iota_{24} & \kappa_{25} \\ \lambda_{31} & \mu_{32} & \nu_{33} & \xi_{34} & o_{35} \\ \pi_{41} & \rho_{42} & \sigma_{43} & \tau_{44} & \upsilon_{45} \\ \varphi_{51} & \chi_{52} & \psi_{53} & \omega_{54} & ☺ \end{pmatrix}$$

ICEF MOSCOW 2011

# Contents

# Preface

This study guide is based on the lecture notes taken by a group of students in my Linear Algebra class at ICEF in 2009. Several handwritten lecture notes were compiled into one big piece with the addition of numerical exercises and extended explanation to the parts of lecture notes that were not explained clearly in the class. Numerical exercises are listed at the end of each chapter, and their main purpose is to help you master your algebraic skills as well as achieve enough confidence that you worked out each individual topic well enough. The pieces missing from the original lecture notes were, in most cases, theorems' proofs, which I had to skip mostly due to time constraints, and which are now included here. They are still essential for students who are used to study mathematics with proofs. Additionally, I included six full-length practice exams (three midterm exams and three final exams) to give you an idea of what will be in the the end of the course and also to help you get prepared and pass these important academic checkpoints.

It is very important for you to keep in mind that this study guide is not a replacement for the attendance of lectures and seminars, but rather a manual containing a detailed inventory of all class topics, which you can use in case if you missed something in the class or simply if you need an extra piece of explanation to the topic. This manual was not designed to be a textbook, and you should not use it as a textbook. But you can use it as lecture notes (in its original sense) and you can also take advantage of the additional problems to check your performance. Each chapter contains a number of "typical" problems which occur frequently in exams. These problems are discussed in detail in examples, which are followed by a plenty of exercises. I strongly recommended that every student go through these exercises before each exam.

Some students complained on the technical nature of problems in this study guide and on the lack of economically-oriented examples. Here I have to remind that the class of Linear Algebra was designed to be an instrumental complement to the other quantitative subjects studied at ICEF. In this light, and also taking into account the very strict time frame we have for teaching this class, obtaining numerical proficiency becomes our first priority, at least for the range of topics we discuss here. Since we teach you very basic matrix techniques, the numerical details of these techniques is what actually matters, and that's why we ask you to work without any aid of computers at home and on exams, and that's why it becomes very technical. As to economic applications and examples, the most of them are related to optimization problems and differential equations, which are the objectives of study in other classes, ones which the class of Linear Algebra is complementing. I thus refer you to the other quantitative subjects since they have better opportunity to provide you with entertaining and relevant economic examples.

I would like to acknowledge the efforts of all students who took lecture notes, helped me to typeset this study guide in LaTeX, and participated in hunting for typos. I decided not to publish the list of names here because it either has to be too long to include everyone, or be unfair. Once again, I thank everyone who contributed to this manual and hope that ICEF readership will appreciate our work. However, even after extensive checks, some typos might be skip our collective attention, and I encourage you to send your comments, suggestions, or simply questions to the lecturer of the class by email.

D.D. Pervouchine

December 4, 2011

# 1 Systems of linear equations

## 1.1 A few introductory examples

Many problems in mathematics stem from the technique for solving systems of simultaneous linear equations. This technique, called Gauss elimination, is one of the most frequently used methods in Linear Algebra.

**Example 1.1.** Find the set of solutions to the following system of linear equations

$$\begin{cases} 2x + y = 3 \\ x + 3y = 4 \end{cases}.$$

**Solution.** By expressing $x$ from the first equation and substituting into the second equation, we get

$$x = 4 - 3y$$

$$2(4 - 3y) + y = 3$$

$$8 - 6y - 6 = 3$$

$$y = 1,\ x = 1.$$

Essentially, we solved this problem by using equivalent transformations of linear equations. □

**Example 1.2.** Solve the following system of linear equations.

$$\begin{cases} x + 2y + 3z = 6 \\ 2x + y + z = 4 \\ x - y + z = 1 \end{cases}$$

**Solution.** Again, we express $x$ and substitute into the second and the third equation.

$$x = 6 - 2y - 3z,$$

$$\begin{cases} 2(6 - 2y - 3z) + y + z = 4 \\ 6 - 2y - 3z - y + z = 1 \end{cases},$$

$$\begin{cases} 12 - 3y - 5z = 4 \\ 6 - 3y - 2z = 1 \end{cases},$$

$$\begin{cases} 3y + 5z = 8 \\ 3y + 2z = 5 \end{cases},$$

$$3z = 3,\ z = 1,\ x = 1,\ y = 1.$$

□

One could notice that all these operations apply only to the coefficients at $x$, $y$, and $z$. We thus can save some paper and ink by writing the system of the linear equations (SLE) in the following *extended matrix* form

$$\left(\begin{array}{rrr|r} 1 & 2 & 3 & 6 \\ 2 & 1 & 1 & 4 \\ 1 & -1 & 1 & 1 \end{array}\right),$$

in which the coefficients at $x$, $y$, and $z$ form a square table, with the free terms being attached to the right after a vertical bar.

In principle, matrices are rectangular tables of numbers. From now on, we reserve the capital Latin letters (typically, from the beginning of the alphabet) to be matrix names, such as

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

while the corresponding small letters denote matrix elements. The same notation is abbreviated by $A = (a_{ij})$, where $i$ and $j$ are running indices which denote the position of the element $a_{ij}$ in the matrix ($i$ and $j$ are not equal to any number). The element $a_{ij}$ is located in the intersection of the $i^{\text{th}}$ row and the $j^{\text{th}}$ column.

## 1.2 Transformations of matrix rows

In what follows we will use the following types of transformations which are applied to matrix rows.

I. The $j^{\text{th}}$ row is multiplied by some constant $c$ and added to the $i^{\text{th}}$ row.

II. The $i^{\text{th}}$ row is multiplied by a non-zero constant $c$.

III. The $i^{\text{th}}$ row and the $j^{\text{th}}$ row are interchanged.

Type I transformations will be denoted by the expression $[i] \mapsto [i] + [j] \times c$. For instance, $[2] \mapsto [2] - [1] \times 2$ means that the first row, multiplied by 2, is subtracted from the second row[1]. Note that in every Type I transformation only the $i^{\text{th}}$ row is affected, while the $j^{\text{th}}$ row is unchanged. For instance, in $[2] \mapsto [2] - [1] \times 2$, the first row stays unchanged.

Type II transformations will be denoted by $[i] \mapsto [i] \times c$. For instance, $[2] \mapsto [2] \times (-1)$ means that the second row is multiplied by $-1$. Multiplication of a row by a constant means that every element of that row is multiplied by the constant.

Similarly, Type III transformations are denoted by $[i] \leftrightarrow [j]$. The three types of transformations and their application to systems of linear equations will be exemplified further in the next section.

[1]Recall that the addition or subtraction of two rows means that their components sum up subtract at the respective positions.

## 1.3 Gauss elimination

The Gauss elimination algorithm is used very frequently for solving systems of linear equations and other related problems. The strategy of Gauss elimination is to create zeroes under the matrix's main diagonal[2], thereby transforming the matrix into so called *upper-triangular* form. The elements below the diagonal of an upper-triangular matrix must be equal to zero. Similarly, all elements above the diagonal of a *lower-triangular* matrix must be equal to zero. The matrix which is both upper-triangular and lower-triangular is called *diagonal* matrix. All elements outside of the diagonal of a diagonal matrix are equal to zero. The n×n diagonal matrix with the numbers $\lambda_1, \ldots, \lambda_n$ on the diagonal will be denoted by $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. The upper-triangular matrix is transformed further to obtain the vector of solutions, as will be shown below.

The essential steps of Gauss elimination are illustrated by the following examples. Note that the elementary transformations described in section 1.2 do not change the set of solutions to the system of linear equations and, thus, create equivalent system of equations at each step.

**Example 1.3.** Solve the following system of linear equations (same as in example 1.2)

$$\begin{cases} x + 2y + 3z = 6 \\ 2x + y + z = 4 \\ x - y + z = 1 \end{cases}.$$

**Solution.**

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 2 & 1 & 1 & 4 \\ 1 & -1 & 1 & 1 \end{array}\right) \xrightarrow{[2]\mapsto[2]-[1]*2} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -3 & -5 & -8 \\ 1 & -1 & 1 & 1 \end{array}\right) \xrightarrow{[2]\mapsto[2]*(-1)}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 3 & 5 & 8 \\ 1 & -1 & 1 & 1 \end{array}\right) \xrightarrow{[3]\mapsto[3]-[1]} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 3 & 5 & 8 \\ 0 & -3 & -2 & -5 \end{array}\right) \xrightarrow{[3]\mapsto[3]+[2]}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 3 & 5 & 8 \\ 0 & 0 & 3 & 3 \end{array}\right) \xrightarrow{[3]\mapsto[3]/3} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 3 & 5 & 8 \\ 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow{[1]\mapsto[1]-[3]*3}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 3 & 5 & 8 \\ 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow{[2]\mapsto[2]-[3]*5} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 3 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow{[2]\mapsto[2]/3}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow{[1]\mapsto[1]-[2]*2} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

□

[2]The main diagonal of a matrix goes south-east from the upper left corner.

That is, the strategy of Gauss elimination is the following.

- Look at the first column and, if necessary, use a type III transformation to make the left-upper-corner element be non-zero. This non-zero element will be called *leading* or *pivot* element. Proceed to the second column if all elements of the first column are equal to zero.

- For each of the rows below the row containing the leading element, use type I transformations with suitable constants to make all elements below the pivot element be equal to zero.

- Proceed to the second column and, if necessary, use a type III transformation to make the element in the position $(2, 2)$ be non-zero. It will now be next pivot element. If all the elements of the second column below the pivot element are equal to zero, proceed to the next column.

- Repeat these steps with other columns until the matrix is in the upper-triangular form.

When transformed to the upper-triangular form, the matrix of a system of linear equations does not immediately give you the set of solutions. For that we need *backward steps* of Gauss elimination, which can be applied to the matrix with non-zero elements of the diagonal.

- Since we assumed that all the elements on the diagonal are non-zero, the element in the position $(n, n)$ is also non-zero. Therefore, a type II transformation can make it equal to one. Then, we can apply type I transformations with suitable $\lambda$'s to make all elements above the pivot element be equal to zero.

- Similarly, we can make all the elements outside of the diagonal be equal to zero, and all elements on the diagonal be equal to one. Then, the set of solutions appears in the column of free terms. Note that this strategy works only when all the diagonal elements after the forward Gauss elimination were non-zero.

The following example shows that the pivot element does not have to be always equal to one, although it is very convenient not to deal with fractions in Gauss elimination.

**Example 1.4.** Solve the following system of linear equations

$$\begin{cases} 2x + 3y + z = 3 \\ 3x - y + 2z = 5 \\ 2x + 2y + z = 3 \end{cases} .$$

**Solution.**

$$\left(\begin{array}{rrr|r} 2 & 3 & 1 & 3 \\ 3 & -1 & 2 & 5 \\ 2 & 2 & 1 & 3 \end{array}\right) \xrightarrow{[2]\mapsto[2]-[1]*\frac{3}{2}} \left(\begin{array}{rrr|r} 2 & 3 & 1 & 3 \\ 0 & -\frac{11}{2} & \frac{1}{2} & \frac{1}{2} \\ 2 & 2 & 1 & 3 \end{array}\right) \xrightarrow{[3]\mapsto[3]-[1]}$$

$$
\left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 0 & -\frac{11}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & -1 & 0 & 0 \end{array}\right) \xrightarrow{[2]\leftrightarrow[3]} \left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 0 & -1 & 0 & 0 \\ 0 & -\frac{11}{2} & \frac{1}{2} & \frac{1}{2} \end{array}\right) \xrightarrow{[2]\mapsto[2]*(-1)}
$$

$$
\left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & -\frac{11}{2} & \frac{1}{2} & \frac{1}{2} \end{array}\right) \xrightarrow{[3]\mapsto[3]+[2]*\frac{11}{2}} \left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{array}\right) \xrightarrow{[3\mapsto[3]*2}
$$

$$
\left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow{[1]\mapsto[1]-[3]} \left(\begin{array}{ccc|c} 2 & 3 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow{[1]\mapsto[1]-[2]*3}
$$

$$
\left(\begin{array}{ccc|c} 2 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow{[1]\mapsto[1]/2} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right).
$$

□

## 1.4 Geometric interpretation of SLE

The geometry standing behind the systems of linear equations is very simple and natural. Consider the system in example 1.1. The first equation, $2x + y = 3$, defines a line in xy-plane. The second equation, $x + 3y = 4$, defines another line. The set of solutions to the SLE is the set of points in xy-plane satisfying both equations, i.e., the intersection of the two lines. In principle, there are three options in the mutual positioning of two lines: intersecting (generic case), parallel, or coinciding. Thus a SLE with two variables may have either one, or zero, or infinitely many solutions. Similarly, the set of solutions to the SLE with three variables shown in example 1.3 is the intersection of three planes in 3-D space. The generic case is again when three planes intersect by one point (the intersection of two planes is a line, and that line intersects the third plane at one point), but there are cases with infinitely many solutions (the three planes have a common line or coincide) or no solutions (parallel planes or lines). Some of these arrangements are shown in Figure 1.

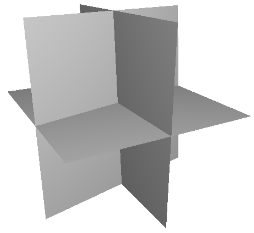
Repeating the same argument for arbitrary number of variables (or reasoning by induction), we arrive to the following conclusion.

**Theorem 1.** *A system of linear equations can have either no solutions, or just one solution, or infinite number of solutions.*

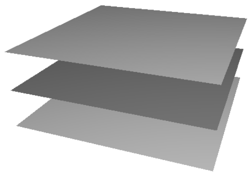
The proof of this statement is not very difficult. However, we will need to introduce the notion of a linear vector space to talk more generally about the set of solutions to an arbitrary system of linear equations. We will do it in the next section.

In both example 1.3 and example 1.4 we dealt with SLE with the unique solution, and thus the final matrix of such SLE was diagonal. This is not true for degenerate systems (i.e., with no or infinitely many solutions). However, Gauss elimination is also applicable to degenerate systems.

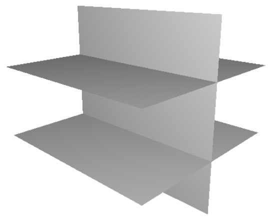
(a) $\infty$ solutions



(b) only 1 solution



(c) no solutions



(d) also no solutions

Figure 1: The intersection of three planes in 3D may consist of zero, one, or infinitely many points.

**Example 1.5.** Solve the following system of linear equations

$$\begin{cases} 2x + 3y + z = 3 \\ 2x + 3y + 2z = 6 \end{cases}.$$

**Solution.**

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 2 & 3 & 2 & 6 \end{array}\right) \xrightarrow{[2]\mapsto[2]-[1]} \left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 0 & 0 & 1 & 3 \end{array}\right) \xrightarrow{[1]\mapsto[1]-[2]} \left(\begin{array}{ccc|c} 2 & 3 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right).$$

We have $x_3 = 3$ and $x_1 = -\frac{3}{2}x_2$. Here $x_2$ can be considered as a parameter, and $x_1$ can be found when $x_2$ is known. $\square$

Note that at the second step the matrix is also upper-triangular, but this time one "step" of the "ladder" is longer than in the former examples. This type of representation, in which the first non-zero element of each row appears to the right and down of the first non-zero element of the previous row, is called matrix's *echelon form*.

## 1.5 Exercises

Problem 1.1. Solve the following systems of linear equations.

(a) $\begin{cases} -3x_1 + 3x_3 = -3 \\ -x_1 + 2x_2 + 2x_4 = -8 \\ -2x_2 + 2x_3 - 2x_4 = 8 \\ 2x_1 + 4x_2 - 2x_3 + 2x_4 = 0 \end{cases}$

(b) $\begin{cases} -5x_1 - 2x_2 + 4x_3 + 3x_4 = -14 \\ -3x_1 - 4x_2 - 4x_3 = 10 \\ -5x_1 + 4x_2 + 4x_3 - 4x_4 = -18 \\ -2x_2 + 3x_3 + x_4 = 1 \end{cases}$

(c) $\begin{cases} 4x_1 - x_2 + 2x_3 - 3x_4 - x_5 = 10 \\ -x_2 + 2x_3 - 3x_4 + x_5 = 10 \\ -2x_1 - 5x_2 + 2x_3 - 4x_4 + 2x_5 = 14 \\ 3x_1 + 4x_2 + 2x_3 - x_4 + 4x_5 = -16 \\ x_1 + 4x_2 - 5x_3 - x_4 - 4x_5 = 13 \end{cases}$

(d) $\begin{cases} 4x_1 + 3x_2 - 4x_3 + 3x_4 = -3 \\ -4x_1 - 3x_2 + 3x_4 = 21 \\ -x_1 + 2x_2 + 4x_3 + 3x_4 = 1 \\ -5x_1 - 4x_2 - x_3 + 4x_4 = 28 \end{cases}$

(e) $\begin{cases} 3x_1 - 5x_2 + 3x_3 + 4x_4 = -27 \\ 4x_1 - 3x_2 - 2x_3 - 3x_4 = 0 \\ 4x_1 - 4x_2 + 3x_4 = -18 \\ -x_1 - 3x_4 = 7 \end{cases}$

(f) $\begin{cases} -3x_1 + 3x_3 - x_4 = -3 \\ 3x_1 - 2x_2 - 2x_3 + 4x_4 = 20 \\ 4x_1 + 4x_2 + 2x_3 + 3x_4 = 25 \\ -4x_1 + 2x_2 - 4x_3 - x_4 = -39 \end{cases}$

(g) $\begin{cases} x_1 - 2x_2 + 2x_3 - 5x_4 = 4 \\ 9x_1 + 2x_2 - 10x_3 - 6x_4 = 10 \\ 3x_1 + 5x_2 - 3x_3 + 3x_4 = 11 \\ 11x_1 + 5x_2 + 10x_4 = 0 \end{cases}$

(h) $\begin{cases} 3x_1 + 11x_2 + 7x_3 - 9x_4 = -1 \\ 12x_1 + 10x_2 - x_3 + 4x_4 = -8 \\ 12x_1 + 3x_2 - x_4 = 3 \\ 4x_1 + 2x_2 + x_3 + 2x_4 = -10 \end{cases}$

(i) $\begin{cases} 5x_1 + 10x_2 + 11x_3 + 9x_4 = -10 \\ 5x_1 - 4x_2 + 12x_3 + 6x_4 = 3 \\ 3x_1 + 5x_2 - 3x_3 + 3x_4 = -6 \\ 6x_1 + 3x_2 - 5x_3 + 5x_4 = -7 \end{cases}$

(j) $\begin{cases} x_1 + 12x_3 - 2x_4 = -1 \\ 5x_1 - 8x_2 + 10x_3 + x_4 = 8 \\ x_1 - 4x_2 - 4x_3 + 10x_4 = -1 \\ 10x_1 - 8x_2 - 8x_3 - x_4 = -5 \end{cases}$

# 2 Linear vector space

## 2.1 Cartesian product of sets

Let $A$ and $B$ be two sets (not necessarily sets of numbers). We define the union $A \cup B$, intersection $A \cap B$, difference $A \backslash B$, and symmetric difference $A \triangle B$ of $A$ and $B$ to be

$$A \cup B = \{x \mid x \in A \,\mathrm{or}\, x \in B\},$$

$$A \cap B = \{x \mid x \in A \,\mathrm{and}\, x \in B\},$$

$$A \backslash B = \{x \mid x \in A \,\mathrm{and}\, x \notin B\},$$

$$A \triangle B = (A \backslash B) \cup (B \backslash A).$$

Additionally, we define another operation on elementary sets called Cartesian product, which describes the set that is composed of all *ordered pairs* of elements of $A$ and $B$:

$$A \times B = \{(x, y) \mid x \in A,\, y \in B\}.$$

For instance, if $A = B = [0, 1]$ then $A \times B$ can be thought of as a square in $xy$-plane formed by the pairs $(x, y)$ such that $0 \leq x, y \leq 1$. If $A = \mathbb{R}$ and $B = [0, 1]$ then $A \times B$ is a horizontal stripe of width 1 in y-direction. If $A = \mathbb{R}$ and $B = \mathbb{Z}$ then $A \times B$ is a grid of horizontal lines crossing $y$-axis at every integer number. Following this definition, the set formed by $n$-tuples of real numbers, $\{(x_1, \ldots, x_n) \mid \forall i : x_i \in \mathbb{R}\}$, is exactly the same set as $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$. Previously, we did not define functions assuming that a function is a "genuine" object which doesn't need a definition. In fact, one can use Cartesian product to give a rigorous definition of a function. This is achieved in two steps.

First, let us assume we are given two sets, $X$ and $Y$, called the *domain set* and the *image set*, respectively. We define a *relation* $F$ on $X$ and $Y$ to be a generic subset of $X \times Y$. For instance, the set $B_2(1) = \{(x, y) \mid x^2 + y^2 \leq 1\}$, which corresponds to a circle on the plane, can be considered to be a relation. However, it differs from our conventional understanding of a function because for each given $x$ such that $(x, y) \in F$, $y$ can take more than one value. Now we will define a function or, more generally, a mapping[3], to be a relation $F$ on $X$ and $Y$ such that $(x, y_1) \in F$ and $(x, y_2) \in F$ necessarily implies $y_1 = y_2$. Restricting this construction on a subset of $X$, if necessary, we can also assume that for every $x \in X$ there exists $y \in Y$ such that $(x, y) \in F$. This fact is denoted by $F : X \to Y$, which spells as "the function $F$ maps the set $X$ to the set $Y$".

The definition of a function ensures that there is one and only one $y$ for every $x$ such that $(x, y) \in F$. As you know, this $y$ is denoted by $y = F(x)$. It is a link from the formal set-theoretic construct, which defines a function to be a special class of relations, whereas in our colloquial understanding a function is the rule, which appoints to every value of $x$ from the domain one and only one value of $y$ from the image. Functions can be of different nature, and the same set-theoretic formalism also allows defining functions of several variables. For instance, a function of two variables, $x_1$ and $x_2$, both from the set $X$, and taking values in the set $Y$ can be thought of as $F : X \times X \to Y$. We need functions to give a formal definition of a linear vector space.

## 2.2 Linear vector space

**Definition 1.** A *vector space over the field of real numbers* consists of three objects: (i) a set $V$ called the *ground set*, (ii) the mapping $f : V \times V \to V$, which is a *binary operation*, and (iii) the mapping $g : \mathbb{R} \times V \to V$. The mapping $f$ is called *addition*; the mapping $g$ is called *multiplication by scalars*. The mapping $f(x, y)$ is usually denoted by the $+$ sign, i.e., $f(x, y) = x + y$, while $x$ and $y$ could be objects which you don't normally add to each other (for instance, students or teachers). Similarly, the mapping $g(\lambda, x)$ is often denoted by $g(\lambda, x) = \lambda x$, although it might have nothing to do with the conventional multiplication. The elements of $V$ will be referred to as *vectors* although they might not be geometric vectors. The object $(V, +, *)$ with the attached operations must obey the following axioms.

1. $\forall x, y \in V : x + y = y + x$
2. $\forall x, y, z \in V : (x + y) + z = x + (y + z)$
3. $\exists 0 \in V : \forall x \in V : x + 0 = 0 + x = x$
4. $\forall x \in V \, \exists (-x) \in V : x + (-x) = (-x) + x = 0$

[3]Speaking rigorously, the term "function" means a mapping which takes numeric values, while for general mappings the image can be any set.

5. $\forall x, y \in V, \forall \lambda \in \mathbb{R} : \lambda(x + y) = \lambda x + \lambda y$

6. $\forall x \in V, \forall \lambda, \mu \in \mathbb{R} : \lambda(\mu x) = (\lambda\mu)x$

The following examples can give you the idea which sets do form linear vector spaces, and which sets do not.

1. The set $\mathbb{R}$ forms a linear vector space with respect to conventional addition of numbers and multiplication by scalars.

2. The set $\mathbb{R}^2 = \{(x, y)\}$ forms a linear vector space with respect to the component-wise addition and multiplication by scalars, i.e., $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$ and $\lambda(x_1, x_2) = (\lambda x_1, \lambda x_2)$.

3. Similarly, the set $\mathbb{R}^n$ forms a linear vector space with respect to component-wise addition and multiplication by scalars, i.e., $(x_1, \dots, x_n)+(y_1, \dots, y_n) = (x_1+y_1, x_2+y_2, \dots, x_n+y_n)$ and $\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$;

4. The set $P_{deg=n}$ of $n$-th degree polynomials is not a linear vector space because $0 \notin P_{deg=n}$.

5. The set $P_{deg\leq n}$ of all polynomials with the degree not greater than $n$ does form a linear vector space

6. The set $C[a, b]$ of all continuous functions defined on $[a, b]$ does form a linear vector space (the sum of continuous functions is again a continuous function).

7. The set $C^n[a, b]$ of all real functions on the segment $[a, b]$ which are continuous up to their $n$-th derivative (in particular, $n$ could be $\infty$) forms a linear vector space.

Note that in most cases the operation comes logically from the naturally existing addition and multiplication by scalars, while being or not being a linear vector space depends on whether or not these operations *stay within the same set* $V$, i.e., whether or not $f(x, y)$ still belongs to $V$. This motivates the following definition.

**Definition 2.** Let $(V, +, *)$ the linear vector space. The subset[4] $U \subseteq V$ is called *a subspace* of $V$ if $(U, +, *)$ forms a linear vector space with respect to the operations $+$ and $*$ that are inherited from $V$ (by restriction on the smaller domain $U$). In other words, $U$ is a subspace of $V$ if $x, y \in U \implies x + y \in U$, $x \in U \implies (-x) \in U$, and $\lambda \in \mathbb{R}$, $x \in U \implies \lambda x \in U$.

Examples of subspaces of linear vector spaces are given below.

1. $\mathbb{R} \subset \mathbb{R}^2$ can be considered as a linear subspace.

2. $P_{deg\leq n} \subset C(-\infty, +\infty)$ is a linear subspace.

3. $C^{n+1}[a, b] \subset C^n[a, b]$ is a linear subspace.

4. The set $B_2(1)$ defined on page 14 is not a linear subspace of $\mathbb{R}^2$.

[4]Please pay attention: $U \subseteq V$ means that $U$ is a subset of $V$, while $U \subset V$ means that $U$ is a *proper* subset of $V$, i.e., $U \subseteq V$ but $U \neq V$. Note that the difference is similar to "$\leq$" and "$<$".

## 2.3 Linear independence

In this section we introduce the notion of linear dependence and linear independence, which are critical to understand the rest of the course.

**Definition 3.** Let $V$ be a linear vector space over the set of real numbers and let $x_1, x_2, \dots, x_n$ be its vectors. Let also $\lambda_1, \lambda_2, \dots, \lambda_n$ be a set of real numbers. The expression $\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n$ is called *linear combination* of vectors $x_1, x_2, \dots, x_n$ with factors $\lambda_1, \lambda_2, \dots, \lambda_n$. The linear combination is called *trivial* if $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$.

**Definition 4.** Let $x_1, x_2, \dots, x_n$ be a set of vectors in a linear space $V$. The set $x_1, x_2, \dots, x_n$ is called *linearly dependent* if there exist numbers $\lambda_1, \lambda_2, \dots, \lambda_n$, not all equal to zero, such that $\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n = 0$. In other words, the set $x_1, x_2, \dots, x_n$ is linearly dependent if some non-trivial linear combination of $x_1, x_2, \dots, x_n$ is equal to zero. The set of vectors that is not linearly dependent is called *linearly independent*.

**Example 2.1.** Consider the following three vectors, $x_1 = (1, 2, 3)$, $x_2 = (0, 1, 2)$, $x_3 = (1, 4, 7)$. We need to check whether or not they are linearly independent. They are linearly dependent if there exist numbers $\lambda_1$, $\lambda_2$, $\lambda_3$ such that $\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 = 0$. That is, $(\lambda_1, 2\lambda_1, 3\lambda_1) + (0, \lambda_2, \lambda_2) + (\lambda_3, 4\lambda_3, 7\lambda_3) = (0, 0, 0)$. This happens if and only if

$$\begin{cases} \lambda_1 + \lambda_3 = 0 \\ 2\lambda_1 + \lambda_2 + 4\lambda_3 = 0 \\ 3\lambda_1 + 2\lambda_2 + 7\lambda_3 = 0 \end{cases}.$$

Solving this SLE by Gauss elimination, we get

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 2 & 1 & 4 & 0 \\ 3 & 2 & 7 & 0 \end{array}\right) \xrightarrow{[2]\mapsto[2]-[1]*2} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 3 & 2 & 7 & 0 \end{array}\right) \xrightarrow{[3]\mapsto[3]-[1]*3}$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 4 & 0 \end{array}\right) \xrightarrow{[3]\mapsto[3]*\frac{1}{2}} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right) \xrightarrow{[3]\mapsto[3]-[2]}$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

As you can see, the third equation degenerates to the equation $0 = 0$, i.e., effectively we have only two equations. This implies that SLE has infinitely many solutions. Hence, there exists at least one non-trivial solution. Thus, we conclude that $x_1$, $x_2$, $x_3$ is a linearly dependent set of vectors.

**Definition 5.** We say that the vector $x$ is expressed linearly through the vectors $x_1, x_2, \dots, x_n$ if $x = \lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n$ for some $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$. In other words, $x$ is expressed linearly through $x_1, x_2, \dots, x_n$ if $x$ is equal to a linear combination of these vectors.

**Theorem 2.** *The set $x_1, x_2, \ldots, x_n$ is linearly dependent if and only if one of $x_1, x_2, \ldots, x_n$ is expressed linearly through the combination of other vectors.*

*Proof.* Assume that $x_i$ is expressed linearly through $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$. That is, $x_i = \lambda_1 x_1 + \cdots + \lambda_{i-1} x_{i-1} + \lambda_{i+1} x_{i+1} + \cdots + \lambda_n x_n$. But then it means that $\lambda_1 x_1 + \cdots + \lambda_{i-1} x_{i-1} - x_i + \lambda_{i+1} x_{i+1} + \cdots + \lambda_n x_n = 0$, i.e., at least one non-trivial (the factor at $x_i$ is equal to $-1$) linear combination of $x_1, \ldots, x_n$ is equal to zero.

Conversely, let $\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n = 0$ be a non-trivial linear combination, i.e., $\lambda_i \neq 0$ for some $i$. Then $-\lambda_i x = \lambda_1 x_1 + \cdots + \lambda_{i-1} x_{i-1} + \lambda_{i+1} x_{i+1} + \cdots + \lambda_n x_n$, and thus $x_i = \frac{\lambda_1}{-\lambda_i} x_1 + \cdots + \frac{\lambda_{i-1}}{-\lambda_i} x_{i-1} + \frac{\lambda_{i+1}}{-\lambda_i} x_{i+1} + \cdots + \frac{\lambda_n}{-\lambda_i} x_n$, i.e., $x_i$ is expressed linearly through $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$. □

Let us comment on the properties of linearly dependent vectors.

1. If the set of vectors $x_1, x_2, \ldots, x_{n-1}$ is linearly dependent, then the set $x_1, x_2, \ldots, x_{n-1}, x_n$ is also linearly dependent.
2. If the set of vectors $x_1, x_2, \ldots, x_n$ contains a zero vector then it is linearly dependent.
3. If the set of vectors $x_1, x_2, \ldots, x_n$ is linearly independent, then the set $x_1, x_2, \ldots, x_{n-1}$ also must be linearly independent.

Spend some time with paper and pencil to prove these properties.

## 2.4 Linear hull and bases

**Definition 6.** Consider $x_1, x_2, \ldots, x_n$, a set of vectors in a linear vector space $V$. The set $\mathcal{L}(x_1, x_2, \ldots, x_n) = \{\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n \mid \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}\}$ is called the *linear hull* of $x_1, x_2, \ldots, x_n$. In other words, the linear hull of $x_1, x_2, \ldots, x_n$ is the set of all their possible linear combinations.

**Definition 7.** The set $x_1, x_2, \ldots, x_n$ of vectors in a linear vector space $V$ is said to *generate* $V$ if $\mathcal{L}(x_1, x_2, \ldots, x_n) = V$. In other words, $x_1, x_2, \ldots, x_n$ generates $V$ if every vector of $V$ can be expressed linearly through $x_1, x_2, \ldots, x_n$.

*Remark.* Note that the notion of linear independence defined in the previous section was not related to the enveloping vector space, i.e., the property of linear independence of a set of vectors is their internal business. In contrast, the property to generate $V$ has a lot to do with the enveloping vector space.

**Definition 8.** The set $x_1, x_2, \ldots, x_n$ of vectors of a linear vector space $V$ is said to be *a base* in $V$ if it generates $V$ and is linearly independent. Sometimes a base is also called a *basis*.

**Theorem 3.** *Let the sets $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_m$ be two bases in a linear vector space $V$. Then $n = m$.*

*Proof.* Assume the contrary, i.e., $n \neq m$. Without loss of generality, $n > m$. Then, every vector $x_1, x_2, \dots, x_n$ can be expressed linearly through $y_1, y_2, \dots, y_m$, as $y_1, y_2, \dots, y_m$ is a base. That is,

$$\begin{aligned} x_1 &= c_{11}y_1 + c_{12}y_2 + \cdots + c_{1m}y_m, \\ x_2 &= c_{21}y_1 + c_{22}y_2 + \cdots + c_{2m}y_m, \\ &\dots \\ x_n &= c_{n1}y_1 + c_{n2}y_2 + \cdots + c_{nm}y_m. \end{aligned}$$

Consider now an arbitrary linear combination of $x_1, x_2, \dots, x_n$ that is equal to zero. We get $\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n = \lambda_1(c_{11}y_1 + c_{12}y_2 + \cdots + c_{1m}y_m) + \lambda_2(c_{21}y_1 + c_{22}y_2 + \cdots + c_{2m}y_m) + \cdots + \lambda_n(c_{n1}y_1 + c_{n2}y_2 + \cdots + c_{nm}y_m) = (c_{11}\lambda_1 + c_{21}\lambda_2 + \cdots + c_{n1}\lambda_n)y_1 + (c_{12}\lambda_1 + c_{22}\lambda_2 + \cdots + c_{n2}\lambda_n)y_2 + \cdots + (c_{1m}\lambda_1 + c_{2m}\lambda_2 + \cdots + c_{nm}\lambda_n)y_m = 0$. This condition is equivalent to the following SLE

$$\left\{ \begin{array}{ccccccccc} c_{11}\lambda_1 & + & c_{21}\lambda_2 & + & \dots & + & c_{n1}\lambda_n & = & 0 \\ c_{12}\lambda_1 & + & c_{22}\lambda_2 & + & \dots & + & c_{n2}\lambda_n & = & 0 \\ c_{1m}\lambda_1 & + & c_{2m}\lambda_2 & + & \dots & + & c_{nm}\lambda_n & = & 0 \end{array} \right. ,$$

which must have at least one non-zero solution since $m < n$ by our supposition. This implies that $x_1, x_2, \dots, x_n$ cannot be linearly independent. Thus $m = n$. □

**Corollary 3.1.** Let $x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n = \mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_n x_n$ be two linear decompositions of the same vector $x$ in the base $x_1, x_2, \dots, x_n$. Then $\forall i : \lambda_i = \mu_i$. That is, the coefficients of a linear decomposition of a vector $x$ in a given base are defined uniquely. These coefficients are called *the coordinates* of the vector $x$ in that base. Note that the coordinates depend on the order, in which $x_1, x_2, \dots, x_n$ are written.

**Example 2.2.** Find the coordinates of the vector $x = (1, 2, 3)$ in the base $e_1 = (1, 1, 0)$, $e_2 = (1, 0, 1)$, and $e_3 = (0, 1, 1)$.

**Solution.** From now on we will follow the notation, in which the components of a vector are written in columns, not rows. It helps to save space on the page and also makes sense in terms of the formal framework, in which a row and a column are considered to be a particular case of rectangular matrix. By definition, we get

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda_1 + \lambda_2 \\ \lambda_1 + \lambda_3 \\ \lambda_2 + \lambda_3 \end{pmatrix}.$$

We thus have the following SLE

$$\left\{ \begin{array}{l} \lambda_1 + \lambda_2 = 1 \\ \lambda_1 + \lambda_3 = 2 \\ \lambda_2 + \lambda_3 = 3 \end{array} \right. ,$$

which is solved by Gauss elimination

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

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right).$$

Thus, the coordinates of $x$ in $e_1$, $e_2$, $e_3$ are $(0, 1, 2)$. Indeed, $x = 0 \cdot e_1 + 1 \cdot e_2 + 2 \cdot e_3$.

**Example 2.3.** Find the coordinates of the function $f(x) = \frac{1}{x^2-5x+6}$ considered as a vector in the respective linear vector space with the base $e_1 = \frac{1}{x-2}$ and $e_2 = \frac{1}{x-3}$.

**Solution.**

$$\frac{1}{x^2 - 5x + 6} = \frac{\lambda_1}{x-2} + \frac{\lambda_2}{x-3} = \frac{\lambda_1(x-3) + \lambda_2(x-2)}{x^2 - 5x + 6}.$$

In order for this to hold, the sum of free terms of the polynomial in the right-hand side of the equation must be equal to the free term of the polynomial in the right-hand side of the equation, and the same must hold for the terms containing $x$. That is,

$$\begin{cases} \lambda_1 + \lambda_2 = 0 \\ -3\lambda_1 - 2\lambda_2 = 1 \end{cases}.$$

Thus, we have

$$\left(\begin{array}{cc|c} 1 & 1 & 0 \\ -3 & -2 & 1 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 1 & 1 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & 1 \end{array}\right).$$

Thus, the coordinates of $f(x)$ in $e_1$, $e_2$ are $(-1, 1)$.

## 2.5 Dimension

The notion of dimension also plays a key role in many branches of mathematics. The intuitive meaning of dimension is that the line is 1-dimensional, the plane is 2-dimensional, the space is 3-dimensional. A bit of generalization takes us to a 4-D space if we agree to include time as the fourth dimension. The dimensions five and higher are more difficult to imagine. However, in all cases the quantity standing behind the notion of dimension is the number of independent directions you can go in your vector space.

**Definition 9.** The dimension of a vector space $V$ is the number of vectors in its every base (see theorem 3). The dimension of a vector space $V$ is denoted by $\dim(V)$.

The dimension of a vector space can be either finite or infinite. In this course we deal with only finite-dimensional spaces. For instance, the set of (not more than) quadratic polynomials is three-dimensional because it spans over $\{1, x, x^2\}$. The dimension of $C[a, b]$, in contrast, is infinite because the set $\{1, x, \ldots, x^n\}$ is linearly independent for an arbitrarily large $n$.

**Theorem 4.** *Let $U$ and $V$ be finite-dimensional vector spaces such that $U \subseteq V$. Then, $\dim(U) \leq \dim(V)$ and $\dim(U) = \dim(V)$ if and only if $U = V$.*

*Proof.* By theorem 3, if $\dim(U) = \dim(V)$ then every base $e_1, e_2, \ldots, e_n$ of $U$ is also the base of $V$ . Then, we have $U \subseteq V$ and $V \subseteq U$, i.e., $U = V$. Now consider the case when $U \subseteq V$ and $U \neq V$. Let $e_1, e_2, \ldots, e_n$ be the base of $U$. Since $V \backslash U \neq \emptyset$, there exists at least one $y \in V \backslash U$. At that, $e_1, e_2, \ldots, e_n$ and $y$ is linearly independent because otherwise either $y \in U$ or $e_1, e_2, \ldots, e_n$ are linearly dependent. We now get the linear space $U_1 = \mathcal{L}(e_1, e_2, \ldots, e_n, y)$. Similarly, if $U_1 \neq V$ then there exists $z \in V$ which is not a linear combination of $e_1, e_2, \ldots, e_n$, and $y$, so we now obtain $U_2 = \mathcal{L}(e_1, e_2, \ldots, e_n, y, z)$. For some $k$ we will get $U_k = V$ because $\dim(V)$ is finite and the growing chain of subspaces must end at some point. Thus, the base of $V$ consists of $e_1, e_2, \ldots, e_n$ and some other vectors, i.e., $\dim(U) < \dim(V)$. $\square$

The procedure described in this theorem is known as *complementation* or *completion* of bases. It says that if one vector space is contained in another, then every base of the smaller vector space can be complemented to be the base of the bigger vector space. Note that it doesn't work for infinite-dimensional vector spaces as, for instance, $P_{deg \leq n} \subset C(-\infty, +\infty)$, but their dimensions are both infinite.

## 2.6 Exercises

Problem 2.1. Determine which of the following sets of vectors are linearly independent. If the set is linearly dependent, chose a set of base vectors and express the rest of the vectors through the base chosen.

(a) $a_1 = \begin{pmatrix} 2 \\ 0 \\ -3 \\ 1 \end{pmatrix}, a_2 = \begin{pmatrix} -2 \\ 3 \\ 1 \\ -5 \end{pmatrix}, a_3 = \begin{pmatrix} -3 \\ -5 \\ 4 \\ -1 \end{pmatrix}, a_4 = \begin{pmatrix} 4 \\ 1 \\ 4 \\ 4 \end{pmatrix}$

(b) $a_1 = \begin{pmatrix} 3 \\ 0 \\ -4 \\ -1 \\ -1 \end{pmatrix}, a_2 = \begin{pmatrix} -3 \\ 1 \\ -5 \\ -3 \\ 3 \end{pmatrix}, a_3 = \begin{pmatrix} -1 \\ 1 \\ -4 \\ 1 \\ -1 \end{pmatrix}, a_4 = \begin{pmatrix} -2 \\ 0 \\ -5 \\ 3 \\ -4 \end{pmatrix}$

(c) $a_1 = \begin{pmatrix} -3 \\ -4 \\ 3 \end{pmatrix}, a_2 = \begin{pmatrix} -1 \\ -3 \\ 4 \end{pmatrix}, a_3 = \begin{pmatrix} 4 \\ 3 \\ -3 \end{pmatrix}, a_4 = \begin{pmatrix} -1 \\ 1 \\ 4 \end{pmatrix}$

(d) $a_1 = \begin{pmatrix} 4 \\ 1 \\ 2 \\ -4 \end{pmatrix}, a_2 = \begin{pmatrix} -2 \\ -1 \\ -4 \\ 3 \end{pmatrix}, a_3 = \begin{pmatrix} 3 \\ 4 \\ 2 \\ 2 \end{pmatrix}, a_4 = \begin{pmatrix} -5 \\ 2 \\ 2 \\ -3 \end{pmatrix}, a_5 = \begin{pmatrix} -5 \\ 4 \\ 2 \\ 1 \end{pmatrix}.$

Problem 2.2. Find the coordinates of the vector $x$ in the corresponding bases

(a) $e_1 = \begin{pmatrix} 1 \\ 2 \\ 7 \end{pmatrix}$, $e_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$, $e_3 = \begin{pmatrix} -1 \\ -2 \\ -8 \end{pmatrix}$, $x = \begin{pmatrix} -3 \\ -5 \\ 1 \end{pmatrix}$

(b) $e_1 = \begin{pmatrix} 6 \\ 3 \\ 4 \end{pmatrix}$, $e_2 = \begin{pmatrix} 9 \\ 4 \\ 8 \end{pmatrix}$, $e_3 = \begin{pmatrix} -4 \\ -2 \\ -3 \end{pmatrix}$, $x = \begin{pmatrix} 0 \\ 1 \\ -4 \end{pmatrix}$

(c) $x = \begin{pmatrix} 3 \\ 4 \\ 6 \\ 5 \end{pmatrix}$, $e_1 = \begin{pmatrix} 5 \\ 1 \\ 2 \\ 2 \end{pmatrix}$, $e_2 = \begin{pmatrix} 1 \\ -4 \\ -5 \\ 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 4 \\ 2 \\ -1 \\ -1 \end{pmatrix}$, $e_4 = \begin{pmatrix} -7 \\ -1 \\ -6 \\ -2 \end{pmatrix}$

(d) $x = \begin{pmatrix} -6 \\ 2 \\ 1 \\ -1 \end{pmatrix}$, $e_1 = \begin{pmatrix} 3 \\ 1 \\ 4 \\ 4 \end{pmatrix}$, $e_2 = \begin{pmatrix} 3 \\ 1 \\ 1 \\ -3 \end{pmatrix}$, $e_3 = \begin{pmatrix} 7 \\ -2 \\ 0 \\ -5 \end{pmatrix}$, $e_4 = \begin{pmatrix} 4 \\ -1 \\ 2 \\ -2 \end{pmatrix}$

(e) $x = \begin{pmatrix} 1 \\ 2 \\ -4 \\ 3 \end{pmatrix}$, $e_1 = \begin{pmatrix} 3 \\ 5 \\ 2 \\ 7 \end{pmatrix}$, $e_2 = \begin{pmatrix} -5 \\ -6 \\ 2 \\ -7 \end{pmatrix}$, $e_3 = \begin{pmatrix} 3 \\ 4 \\ 5 \\ 8 \end{pmatrix}$, $e_4 = \begin{pmatrix} 1 \\ 1 \\ -7 \\ -3 \end{pmatrix}$

# 3 Fundamental set of solutions

## 3.1 Homogeneity

We will now give a formal overview of the method for solving arbitrary systems of linear equations. Consider a system of linear equations with $n$ variables and $m$ equations of the following general form.

$$\left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m \end{array} \right. ,$$

which is written in the extended matrix form as

$$\left( \begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right).$$

**Definition 10.** A system of linear equations is called *homogeneous* if $b_1 = b_2 = \cdots = b_n = 0$. Otherwise, it is called *non-homogeneous*. Homogeneous and non-homogeneous systems of linear equations will be abbreviated as HSLE and NHSLE, respectively.

**Theorem 5.** *The set of solutions to a homogeneous system of linear equations is a linear vector space.*

*Proof.* Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ be two solutions to a HSLE. Since $a_{i1}(x_1 + y_1) + \cdots + a_{in}(x_n + y_n) = (a_{i1}x_1 + \cdots + a_{in}x_n) + (a_{i1}y_1 + \cdots + a_{in}y_n) = 0 + 0 = 0$ for all $i$, then $x + y$ is a solution, too. Similarly, $\lambda x = (\lambda x_1, \ldots, \lambda x_n)$ and $0 = (0, \ldots, 0)$ are also solutions to the same SLE. Therefore, the set of solutions to a homogeneous system of linear equations forms a linear vector space. □

We will sometimes refer to a *general solution* to the system of linear equations as an expression describing all possible solutions, as it may have many, while a *particular solution* refers to one particularly chosen solution.

**Theorem 6.** *A general solution to NHSLE can be expressed as a particular solution to NHSLE + general solution to the respective HSLE.*

*Proof.* This statement follows from the observation that if $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ are two solutions to the NHSLE, then $x - y$ is a solution to the HSLE. Indeed, $a_{i1}(x_1 - y_1) + \cdots + a_{in}(x_n - y_n) = (a_{i1}x_1 + \cdots + a_{in}x_n) - (a_{i1}y_1 + \cdots + a_{in}y_n) = b_i - b_i = 0$ for all $i$. Similarly, if $x = (x_1, \ldots, x_n)$ is a solution to NHSLE and $y = (y_1, \ldots, y_n)$ is a solution to HSLE, then $x + y$ is a solution to NHSLE because $a_{i1}(x_1 + y_1) + \cdots + a_{in}(x_n + y_n) = (a_{i1}x_1 + \cdots + a_{in}x_n) + (a_{i1}y_1 + \cdots + a_{in}y_n) = b_i + 0 = b_i$. □

## 3.2 Rank of a matrix

Recall that by theorem 5 in the previous section, the set of solutions to a homogeneous system of linear equations is a linear vector space. The dimension of this vector space is related to the dimension of the vector space generated by the rows of the matrix of SLE.

**Definition 11.** Let $A$ be a m×n matrix. The maximum number of linearly independent rows of $A$ is called the *rank* of matrix $A$ and is denoted by $\mathrm{rk}(A)$.

It follows immediately from this definition and from the Gauss elimination method that the rank of a matrix is equal to the number of non-zero rows in the matrix echelon form (page 12).

**Definition 12.** Let $A$ be a m×n matrix. The *transpose* to $A$ is the matrix $A^T$ which is obtained by writing the columns of $A$ as rows or, equivalently, by writing the rows of $A$ as columns. That is,

$$A^T = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}, \text{ where } A = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ a_{12} & a_{22} & \ldots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{pmatrix}.$$

It turns out that the maximum number of linearly-independent rows and the maximum number of linearly-independent columns are the same. That is, $\mathrm{rk}(A)$ is independent of transposition of rows and columns, i.e., $\mathrm{rk}(A^T) = \mathrm{rk}(A)$.

Consider a set of vectors $a_1 = (a_{11}, a_{12}, \dots, a_{1n})$, $a_2 = (a_{21}, a_{22}, \dots, a_{2n})$, $a_m = (a_{m1}, a_{m2}, \dots, a_{mn})$ in $\mathbb{R}^n$ formed by the rows of matrix $A$. Then, $\mathrm{rk}(A) = dim\mathcal{L}(a_1, a_2, \dots, a_m)$. The following is the fundamental theorem about the rank of a matrix, which defines SLE, and the dimension of the solution space.

**Theorem 7.** *Let $V$ be the linear space of solutions to the HSLE defined by the matrix $A$. Then,* $\dim(V) = n - \mathrm{rk}(A)$.

Before we proceed with the proof of the theorem, let us consider the following example.

**Example 3.1.** Describe the set of solutions of

$$\begin{cases} x_1 + 2x_2 + x_4 = 0 \\ x_1 + x_2 - x_3 = 0 \end{cases}.$$

**Solution.** The Gauss elimination

$$\left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} \underline{1} & 0 & -2 & -1 & 0 \\ 0 & \underline{1} & 1 & 1 & 0 \end{array}\right)$$

brings us to the SLE defined by

$$\begin{cases} x_1 - 2x_3 - x_4 = 0 \\ x_2 + x_3 + x_4 = 0 \end{cases}.$$

Note that the variables $x_3$ and $x_4$ are "free", i.e., they can take any values without restrictions, while $x_1$ and $x_2$ can be expressed through $x_3$ and $x_4$ as the underlined coefficients are not equal to zero. We thus can use the following strategy: give to one of the free variables, $x_3$ or $x_4$, the value of 1, and set the other free variables to zero. Then, $x_1$ and $x_2$ can be found from the matrix of SLE. Namely, the values of $x_1$ and $x_2$ stand in the corresponding column of the matrix, up to multiplication by -1. This is illustrated by the following schema.

$$\begin{array}{ccc} x_1 & 2 & 1 \\ x_2 & -1 & -1 \\ x_3 & 1 & 0 \\ x_4 & 0 & 1 \end{array}.$$

Theorem 7 says that $\dim(V) = n - \mathrm{rk}(A) = 4 - 2 = 2$. On the other hand, the two vectors found above, $e_1 = (2, -1, 1, 0)$ and $e_2 = (1, -1, 0, 1)$ are linearly independent by construction. Therefore, the set of solutions spans on $e_1$ and $e_2$, i.e., every solution to the original system of linear equations can be expressed as a linear combination of $e_1$ and $e_2$, i.e.,

$$x = \lambda_1 \begin{pmatrix} 2 \\ -1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

*Proof of theorem 7.* First, note that elementary transformations of rows do not affect the set of solutions because they give an equivalent SLE's on each step. It is enough to consider the case of under-defined system ($m < n$), as the statement of the theorem is trivial for the defined and over-defined systems. Essentially, the idea of the proof follows the previous example, where we used Gauss elimination to transform the matrix $A$ to the form

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2m} & \dots & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3m} & \dots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mm} & \dots & a_{mn} \end{pmatrix} \to \begin{pmatrix} * & * & \dots & * & * & \dots & * \\ 0 & * & \dots & * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & * & * & \dots & * \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix},$$

if necessary exchanging the order of variables. Then, the dimension of the solution space is exactly the number of free variables, which is equal to the total number of variables minus the number of equations in the transformed matrix. The number of variables is equal to $n$, and the number of equations in the transformed matrix is equal to $\mathrm{rk}(A)$, i.e., $\dim(V) = n - \mathrm{rk}(A)$. □

## 3.3 Fundamental set of solutions

The technique explained in example 3.1 in the previous section gave rise to an algorithm called *fundamental system of solutions*. First, the matrix of a SLE is transformed to the echelon form. At that, the element of the matrix that were used to create zeros in other rows are called *pivoting elements*. The columns which do not contain pivoting elements are declared to correspond to free variables. Then, one by one, we give the value of 1 to each of the free variables and assign zero to the rest of the free variables; the dependent variables are calculated from these data. Taking into account the conclusion of theorem 6, this also allows solving NHSLE, as we show below.

**Example 3.2.** Find the fundamental set of solutions to

$$\begin{cases} x_1 + 2x_2 + x_3 - x_4 + x_5 = 4 \\ x_1 + x_2 - x_3 + 2x_4 - x_5 = 2 \end{cases}.$$

**Solution.**

$$\left(\begin{array}{rrrrr|r} 1 & 2 & 1 & -1 & 1 & 4 \\ 1 & 1 & -1 & 2 & -1 & 2 \end{array}\right) \to \left(\begin{array}{rrrrr|r} 1 & 2 & 1 & -1 & 1 & 4 \\ 0 & 1 & 2 & -3 & 2 & 2 \end{array}\right) \to$$

$$\left(\begin{array}{rrrrr|r} 1 & 0 & -3 & 5 & -3 & 0 \\ 0 & 1 & 2 & -3 & 2 & 2 \end{array}\right)$$

| | $e_1$ | $e_2$ | $e_3$ | $NHSLE$ |
|---|---|---|---|---|
| $x_1$ | 3 | −5 | 3 | 0 |
| $x_2$ | −2 | 3 | −2 | 2 |
| $x_3$ | 1 | 0 | 0 | 0 |
| $x_4$ | 0 | 1 | 0 | 0 |
| $x_5$ | 0 | 0 | 1 | 0 |

First, we set $x_3 = 1$, $x_4 = 0$, and $x_5 = 0$, and compute $x_1$ and $x_2$. The values for $x_1$ and $x_2$ are opposite to the numbers in the third column of the transformed matrix. Then, we set $x_3 = 0$, $x_4 = 1$, and $x_5 = 0$, and compute $x_1$ and $x_2$ again. This time, $x_1$ and $x_2$ are the numbers that are opposite to those in the fourth column. When we need a particular solution to NHSLE, we can set $x_3 = 0$, $x_4 = 0$, and $x_5 = 0$ and compute $x_1$ and $x_2$ again, this time by using the numbers to the right of the bar. Finally,

$$x = \lambda_1 \begin{pmatrix} 3 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -5 \\ 3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 3 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

. □

## 3.4 Exercises

PROBLEM 3.1. Find ranks of the following matrices

(a) $\begin{pmatrix} 4 & 8 & -4 & 16 & -8 \\ -1 & -2 & 1 & -4 & 2 \\ 3 & 6 & -3 & 12 & -6 \\ 2 & 4 & -2 & 8 & -4 \end{pmatrix}$

(b) $\begin{pmatrix} -2 & -10 & -10 & 2 & 4 \\ -2 & -10 & -10 & 2 & 4 \\ 1 & 5 & 5 & -1 & -2 \\ 2 & 10 & 10 & -2 & -4 \end{pmatrix}$

(c) $\begin{pmatrix} 9 & 10 & -16 & 11 & 3 \\ 9 & 8 & -8 & 7 & -1 \\ -22 & -18 & -20 & 15 & 11 \\ 4 & -3 & -1 & 12 & -10 \end{pmatrix}$

(d) $\begin{pmatrix} 12 & -2 & -3 & 12 & -2 \\ 12 & 8 & 3 & 12 & -4 \\ -20 & 0 & 3 & -20 & 4 \\ 12 & -12 & -9 & 12 & 0 \end{pmatrix}$

(e) $\begin{pmatrix} 4 & -8 & -4 & -16 & -12 \\ -4 & 8 & 4 & 16 & 12 \\ -3 & 6 & 3 & 12 & 9 \\ 3 & -6 & -3 & -12 & -9 \end{pmatrix}$

(f) $\begin{pmatrix} 0 & 1 & -3 & -1 & -4 \\ -16 & 16 & 4 & 0 & -8 \\ 16 & -13 & -13 & -3 & -4 \\ -8 & 6 & 8 & 2 & 4 \end{pmatrix}$

(g) $\begin{pmatrix} 14 & -14 & 4 & 22 & 14 \\ -8 & -16 & 20 & -4 & -8 \\ 0 & -14 & 13 & 5 & 0 \\ 6 & -16 & 11 & 13 & 6 \end{pmatrix}$

(h) $\begin{pmatrix} -3 & -3 & 9 & -15 & -9 \\ 3 & 3 & -9 & 15 & 9 \\ 5 & 5 & -15 & 25 & 15 \\ -2 & -2 & 6 & -10 & -6 \end{pmatrix}$

(i) $\begin{pmatrix} 5 & -2 & 4 & -4 & -4 \\ 12 & -9 & 16 & -8 & -14 \\ -1 & -8 & 12 & 4 & -8 \\ -9 & 12 & -20 & 4 & 16 \end{pmatrix}$

(j) $\begin{pmatrix} 10 & -14 & -2 & -6 & -4 \\ -19 & 11 & 5 & 21 & 4 \\ -15 & 8 & 4 & 17 & 3 \\ 13 & -13 & -3 & -11 & -4 \end{pmatrix}$

PROBLEM 3.2. Find the fundamental system of solutions to the following systems of linear equations

(a) $\begin{cases} x_1 - x_2 + 5x_3 - 13x_4 = 3 \\ x_1 + x_2 - x_3 + x_4 = -3 \end{cases}$

(b) $\begin{cases} 3x_1 - 5x_2 - 3x_3 + 13x_4 = 10 \\ 2x_1 - x_2 + 5x_3 + 4x_4 = -5 \end{cases}$

(c) $\begin{cases} 2x_1 - x_2 - 2x_3 - 5x_4 = -7 \\ 2x_1 + x_2 - 6x_3 + x_4 = -1 \end{cases}$

(d) $\begin{cases} 3x_1 + x_2 + x_3 + 6x_5 = -6 \\ 2x_1 + 3x_2 + 2x_3 - 4x_4 + 5x_5 = -4 \\ 3x_1 + 3x_2 + 5x_3 + 8x_4 - 16x_5 = -6 \end{cases}$

(e) $\begin{cases} -2x_2 + x_3 - 7x_4 - 4x_5 = 7 \\ x_1 + 2x_2 - x_3 + x_4 - 2x_5 = -7 \\ 3x_1 + 4x_2 - 2x_3 - 4x_4 - 10x_5 = -14 \end{cases}$

(f) $\begin{cases} x_1 - 6x_2 - 3x_3 - 17x_4 - 17x_5 = -1 \\ 5x_2 + 7x_3 - 8x_4 - 17x_5 = -5 \\ 3x_2 + 5x_3 - 8x_4 - 15x_5 = -3 \end{cases}$

(g) $\begin{cases} 2x_1 + x_2 + 9x_3 - 15x_4 = 9 \\ x_1 + 2x_3 - 6x_4 + x_5 = 1 \end{cases}$

(h) $\begin{cases} x_1 + x_2 + 2x_3 + 2x_4 - 2x_5 = 0 \\ x_2 + 4x_3 + 7x_4 - 5x_5 = -2 \end{cases}$

(i) $\begin{cases} 3x_1 - 5x_2 + 18x_3 + x_4 - 3x_5 = 2 \\ x_1 + x_2 - 2x_3 + 11x_4 - x_5 = 6 \end{cases}$

(j) $\begin{cases} x_1 - x_2 - 3x_3 + 7x_5 = 0 \\ x_1 + 2x_3 - x_4 + 7x_5 = 1 \end{cases}$

# 4 The determinant and its applications

The determinant of a n×n matrix A is a value that can be obtained by multiplying certain sets of entries of A, and adding and subtracting such products, according to the rules that are discussed below.

## 4.1 Special cases 2×2 and 3×3

The determinant $\det(A)$ of a 2×2 matrix $A$ is defined by

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

The vectors $(a_{11}, a_{12})$ and $(a_{21}, a_{22})$ form a parallelogram in the $xy$-plane with vertices at $(0,0)$, $(a_{11}, a_{12})$, $(a_{21}, a_{22})$, and $(a_{11}+a_{21}, a_{21}+a_{22})$. The absolute value of $a_{11}a_{22} - a_{12}a_{21}$ is the area of the parallelogram. That is, the determinant of a 2×2 matrix is somewhat similar to the area. The absolute value of the determinant together with the sign becomes the *oriented area* of the parallelogram. The oriented area is the same as the usual area, except that it changes its sign when the two vectors forming the parallelogram are interchanged.

The determinant $\det(A)$ of a 3×3 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

is defined by $\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{31}a_{22}a_{13} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$.

Similarly to the 2×2 case, the geometric interpretation of the determinant is the volume of a cuboid spanning the vectors $(a_{11}, a_{12}, a_{13})$, $(a_{21}, a_{22}, a_{23})$, and $(a_{31}, a_{32}, a_{33})$. Again, the absolute value of the determinant together with the sign gives the *oriented volume* of the cuboid.

The determinant expression grows rapidly with the size of the matrix (an n×n matrix contributes $n!$ terms). In the next section we give the formal expression for $\det(A)$ for arbitrary size matrices, which subsumes these two cases, while for now we list the properties of the determinant, which actually determine it as a function uniquely, and continue the sequence of term "length-area-volume" to something that can be called $n$-dimensional volume.

**Example 4.1.** $\det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 1 \times 4 - 2 \times 3 = 4 - 6 = -2$

**Example 4.2.** $\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 1\times5\times9+4\times8\times3+2\times6\times7-7\times5\times3-4\times2\times9-1\times8\times6 =$
$45 + 96 + 84 - 105 - 72 - 48 = 0$

## 4.2 Properties of the determinant

Recall the three types of transformations we introduced to operate with matrices of linear equations (page 8). In type I transformation, the $j^{\text{th}}$ row is multiplied by some constant $c$ and added to the $i^{\text{th}}$ row. In type II transformation, the $i^{\text{th}}$ row is multiplied by a non-zero constant $c$. Type III transformation interchanges the $i^{\text{th}}$ row and the $j^{\text{th}}$ row of the matrix. The following are the properties of the determinant.

1. $\det(A)$ doesn't change under type I transformations, i.e., the addition of a row (or column), factored by a number, to another row (or column) doesn't change $\det(A)$.

2. $\det(A)$ is multiplied by $c$ under type II transformations, i.e., if a row or column is factored by a number, then $\det(A)$ is multiplied by the same number.

3. $\det(A)$ is multiplied by -1 under type III transformations, i.e., $\det(A)$ changes sign if two rows or two columns are interchanged.

These three properties hold for the area of a parallelogram and for the volume of a cuboid. They must also hold for higher-dimensional analogs of these measures, and these requirements define $\det(A)$ uniquely up to a constant factor. If we additionally require that $\det(E) = 1$, where

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

is the n×n identity matrix, then with the requirements (1)-(3) above it can be used as a definition of the determinant.

## 4.3 Formal definition

We define a *permutation* $(i_1, i_2, \dots, i_n)$ of the set of integer numbers $(1, 2, \dots, n)$ as a mapping $f : (1, 2, \dots, n) \to (1, 2, \dots, n)$ such that no pair of numbers $i$ and $j$ are mapped to the same number, i.e., $f(i) = f(j)$ implies $i = j$. It means that $i_1, i_2, \dots, i_n$ is the same set of number as $1, 2, \dots, n$, simply written in a different order. For instance, the permutation $(2, 3, 1, 4)$ of numbers $(1, 2, 3, 4)$ is the mapping $f$ such that $f(1) = 2$, $f(2) = 3$, $f(3) = 1$, and $f(4) = 4$, whereas if $f(1) = f(2)$ then $f$ is not a permutation.

Each permutation $(i_1, i_2, \dots, i_n)$ contains *orders* and *inversions*. For instance, $(2, 3, 1, 4)$ contains four orders, $(2, 3)$,$(2, 4)$, $(3, 4)$, and $(1, 4)$ and two inversions, $(2, 1)$ and $(3, 1)$. An inversion happens when a bigger number precedes a smaller number; all other pairs are called orders. We define the sign of a permutation, $\sigma(i_1, i_2, \dots, i_n)$ be equal to $-1$ to the power of the number of inversions. For instance, $\sigma(2, 3, 1, 4) = 1$ and $\sigma(2, 1, 3, 4) = -1$.

The determinant of an arbitrary n×n matrix is defined by

$$\det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \sum_{(i_1, i_2, \dots, i_n)} \sigma(i_1, i_2, \dots, i_n) a_{1i_1} a_{2i_2} \dots a_{ni_n},$$

where the sum runs over all possible permutations of the set $(1, 2, \dots, n)$. Clearly, there are $n! = 1 \cdot 2 \cdot \dots \cdot n$ such permutations and, thus, the general determinant expression contains $n!$ terms. One can also say that the determinant of a matrix $A = (a_{ij})$ is the sum of products of matrix elements such that in every product term one and only one element is taken from each row and column.

**Theorem 8.** *The determinant of an upper-triangular square matrix is equal to the product of its diagonal entries.*

*Proof.* Let $A$ be an upper-triangular matrix, i.e.,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix}.$$

The sum of products of matrix elements such that one and only one element is taken from each row and column always contains an element located below the diagonal, all of which are equal to zero, with the exception of the term which is the product of the diagonal entries, i.e., $a_{11}a_{22}\dots a_{nn}$. This completes the proof. □

## 4.4 Laplace's theorem

The minor $M_{i,j}$ of a matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1\,j-1} & a_{1\,j} & a_{1\,j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1\,1} & \dots & a_{i-1\,j-1} & a_{i-1\,j} & a_{i-1\,j+1} & \dots & a_{i-1\,n} \\ a_{i1} & \dots & a_{i\,j-1} & a_{i\,j} & a_{i\,j+1} & \dots & a_{i\,n} \\ a_{i+1\,1} & \dots & a_{i+1\,j-1} & a_{i+1\,j} & a_{i+1\,j+1} & \dots & a_{i+1\,n} \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{n\,j-1} & a_{n\,j} & a_{n\,j+1} & \dots & a_{nn} \end{pmatrix}$$

is defined to be the determinant of the (n-1)×(n-1) matrix that is obtained from $A$ by removing the i$^{\text{th}}$ row and the j$^{\text{th}}$ column:

$$M_{ij} = \det \begin{pmatrix} a_{11} & \dots & a_{1\,j-1} & a_{1\,j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ a_{i-1\,1} & \dots & a_{i-1\,j-1} & a_{i-1\,j+1} & \dots & a_{i-1\,n} \\ a_{i+1\,1} & \dots & a_{i+1\,j-1} & a_{i+1\,j+1} & \dots & a_{i+1\,n} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{n\,j-1} & a_{n\,j+1} & \dots & a_{nn} \end{pmatrix}$$

The expression $M_{ij}$ is also known as *algebraic complement* to $a_{ij}$. Then the determinant of A is given by the following *Laplace's* formula.

**Theorem 9.**

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij} \text{ (decomposition by the j-th column)}$$

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij} \text{ (decomposition by the i-th row)}$$

The Laplace's theorem gives a useful tool for the calculation of determinants, one in which the n$^{\text{th}}$ order determinant is expressed through lower order minors recursively. Below we go through examples of these recursive calculations.

**Example 4.3.** $\begin{vmatrix} 1 & 0 & 1 & 2 \\ 2 & 0 & 1 & 1 \\ -1 & -1 & 3 & 1 \\ 2 & 0 & 1 & 2 \end{vmatrix} = (-1)^{1+2} \times 0 \times \begin{vmatrix} 2 & 1 & 1 \\ -1 & 3 & 1 \\ 2 & 1 & 2 \end{vmatrix} + (-1)^{2+2} \times 0 \times \begin{vmatrix} 1 & 1 & 2 \\ -1 & 3 & 1 \\ 2 & 1 & 2 \end{vmatrix} +$
$(-1)^{2+3} \times (-1) \times \begin{vmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 2 & 1 & 2 \end{vmatrix} + 0 = 2 + 2 + 4 - 4 - 1 = -1$

This example shows that the Laplace's formula is most useful for decomposition by a column or a row which contains many zeros. If there is no such row or column, you may to create it by using type I transformations as they don't affect the determinant.

**Example 4.4.** $\begin{vmatrix} 1 & 1 & 1 & -1 \\ 2 & 1 & 1 & 2 \\ 3 & -1 & 1 & 1 \\ 5 & 2 & 1 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & -1 \\ 1 & 0 & 0 & 3 \\ 3 & -1 & 1 & 1 \\ 5 & 2 & 1 & -1 \end{vmatrix} = 1 \times (-1)^{2+1} \times$
$\times \begin{vmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 2 & 1 & -1 \end{vmatrix} + 0 \times (-1)^{2+2} \begin{vmatrix} 1 & 1 & -1 \\ 3 & 1 & 1 \\ 5 & 1 & -1 \end{vmatrix} + 0 \times (-1)^{2+3} \begin{vmatrix} 1 & 1 & -1 \\ 3 & -1 & 1 \\ 5 & 2 & -1 \end{vmatrix} + 3 \times (-1)^{2+4} \begin{vmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \\ 5 & 2 & 1 \end{vmatrix} =$
$-2 + 30 = 28.$

## 4.5 Minors and ranks

The notion of a minor can be used to assess the rank of an arbitrary n×m matrix. Consider two sets of indices, $1 \leq i_1 \leq i_2 \leq \cdots \leq i_k \leq n$ and $1 \leq j_1 \leq j_2 \leq \cdots \leq j_k \leq m$, and consider the submatrix of $A$ which consists of the elements located on the intersection of rows $i_1, i_2, \ldots i_k$ and $j_1, j_2, \ldots j_k$:

$$\begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \ldots & a_{i_1 j_k} \\ a_{i_2 j_1} & a_{i_2 j_2} & \ldots & a_{i_2 j_k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_k j_1} & a_{i_k j_2} & \ldots & a_{i_k j_k} \end{pmatrix}.$$

Its determinant is called a $k^{\text{th}}$ *order minor* of the matrix $A$. For instance, if

$$A = \begin{pmatrix} 1 & -1 & -8 & 1 & 2 \\ 0 & 8 & -4 & -5 & 6 \\ -6 & 9 & 0 & 6 & -1 \\ 2 & 1 & 2 & 3 & 6 \end{pmatrix},$$

$(i_1, i_2, i_3) = (1, 2, 4)$, and $(j_1, j_2, j_3) = (2, 3, 5)$, then the corresponding 3$^{\text{rd}}$ order minor is

$$\begin{vmatrix} -1 & -8 & 2 \\ 8 & -4 & 6 \\ 1 & 2 & 6 \end{vmatrix} = 24 - 48 + 32 + 8 + 6 + 384 + 12 = 418.$$

**Theorem 10.** *Let $A$ be a $n \times m$ matrix. If at least one $k^{\text{th}}$ order minor of $A$ is not equal to zero then* $\operatorname{rk}(A) \geq k$. *If all $k^{\text{th}}$ order minors of $A$ are equal to zero then* $\operatorname{rk}(A) < k$.

**Example 4.5.**

$$\operatorname{rk}\begin{pmatrix} 2 & 1 & 0 & 0 & 3 & 7 & -2 \\ 1 & 0 & 1 & 0 & -1 & 2 & -1 \\ -3 & 0 & 0 & 1 & 2 & 3 & 1 \end{pmatrix} = ?$$

**Solution.** The matrix contains the minor $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \neq 0$. Since the identity matrix has rank three, the rank of the original matrix is not smaller than three. Hence, it is equal to three.

## 4.6 Cramer's rule

Consider the general form of a system of $n$ linear equations with $n$ variables

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \dots & + & a_{nn}x_n & = & b_n \end{cases},$$

and denote $\Delta = \det\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$,

$\Delta_1 = \det\begin{pmatrix} b_1 & a_{12} & \dots & a_{1n} \\ b_2 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_n & a_{n2} & \dots & a_{nn} \end{pmatrix}$, $\Delta_2 = \det\begin{pmatrix} a_{11} & b_1 & \dots & a_{1n} \\ a_{21} & b_2 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & b_n & \dots & a_{nn} \end{pmatrix}$, and so on for all columns, where the $b$-column replaces the i$^{\text{th}}$ column in $\Delta_i$.

**Theorem 11** (Cramer's rule)**.** *Let $(x_1, x_2, \dots, x_n)$ be a solution to the system of linear equations given above. If $\Delta \neq 0$ then $(x_1, x_2, \dots, x_n)$ is unique and $x_i = \frac{\Delta_i}{\Delta}$ for all $i$.*

**Example 4.6.** Consider the system of linear equations given by

$$\left(\begin{array}{cc|c} 2 & 3 & 5 \\ 2 & -1 & 1 \end{array}\right)$$

**Solution.**

$$\Delta = \begin{vmatrix} 2 & 3 \\ 2 & -1 \end{vmatrix} = -2 - 6 = -8,$$

$$\Delta_1 = \begin{vmatrix} 5 & 3 \\ 1 & -1 \end{vmatrix} = -5 - 3 = -8,$$

$$\Delta_2 = \begin{vmatrix} 2 & 5 \\ 2 & 1 \end{vmatrix} = 2 - 10 = -8,$$

$x_1 = \frac{\Delta_1}{\Delta} = \frac{-8}{-8} = 1$, $x_2 = \frac{\Delta_2}{\Delta} = \frac{-8}{-8} = 1$.

**Example 4.7.** Consider the system of linear equations given by

$$\begin{cases} x_1 + 2x_2 + x_3 = 3 \\ x_1 - x_2 + 2x_3 = 0 \\ 3x_1 - x_2 + 7x_3 = 2 \end{cases}$$

**Solution.**

$$\left(\begin{array}{rrr|r} 1 & 2 & 1 & 3 \\ 1 & -1 & 2 & 0 \\ 3 & -1 & 7 & 2 \end{array}\right)$$

$$\Delta = \begin{vmatrix} 1 & 2 & 1 \\ 1 & -1 & 2 \\ 3 & -1 & 7 \end{vmatrix} = -7 + 12 - 1 + 4 - 14 + 2 = -5,$$

$$\Delta_1 = \begin{vmatrix} 3 & 2 & 1 \\ 0 & -1 & 2 \\ 2 & -1 & 7 \end{vmatrix} = -21 + 8 + 0 + 2 - 0 + 6 = -11 + 6 = -5,$$

$$\Delta_2 = \begin{vmatrix} 1 & 3 & 1 \\ 1 & 0 & 2 \\ 3 & 2 & 7 \end{vmatrix} = 0 + 18 + 2 - 0 - 21 - 4 = -5,$$

$$\Delta_3 = \begin{vmatrix} 1 & 2 & 3 \\ 1 & -1 & 0 \\ 3 & -1 & 2 \end{vmatrix} = -2 + 0 - 3 + 9 - 4 - 0 = 0,$$

$$x_1 = \frac{\Delta_1}{\Delta} = \frac{-5}{-5} = 1,$$

$$x_2 = \frac{\Delta_2}{\Delta} = \frac{-5}{-5} = 1,$$

$$x_3 = \frac{\Delta_3}{\Delta} = \frac{0}{-5} = 0.$$

An extension of the Cramer's rule is sometimes used to distinguish between the case of no solutions and the case of infinitely many solutions when the system of linear equations is degenerate.

**Example 4.8.** $\begin{cases} x_1 + x_2 = 2 \\ x_1 + x_2 = 3 \end{cases}$

**Solution.** This system of linear equations has no solutions since $2 \neq 3$. Look what happens when we apply Cramer's rule.

$$\left(\begin{array}{cc|c} 1 & 1 & 2 \\ 1 & 1 & 3 \end{array}\right)$$

$$\Delta = 0,$$

$$\Delta_1 = \begin{vmatrix} 2 & 1 \\ 3 & 1 \end{vmatrix} \neq 0.$$

This situation corresponds to having two parallel lines, $x_1 + x_2 = 2$ and $x_1 + x_2 = 3$, which "intersect" at $x = \infty$, and thus the Cramer's rule gives $\Delta_1/\Delta = \Delta_1/0 = \infty$.

**Example 4.9.** $\begin{cases} x_1 + x_2 = 1 \\ 2x_1 + 2x_2 = 2 \end{cases}$

**Solution.** This system of linear equations has infinitely many solutions since $x_1 + x_2 = 1$ and $2x_1 + 2x_2 = 2$ have the same set of solutions. Then,

$$\left(\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & 2 & 2 \end{array}\right)$$

$$\Delta = 0,$$

$$\Delta_1 = \Delta_2 = \begin{vmatrix} 1 & 1 \\ 2 & 2 \end{vmatrix} = 0$$

Thus, we have infinitely many solutions, as hinted by the formal application of Cramer's rule, which gives "$x = 0/0$".

**Theorem 12** (Kronecker-Capelli)**.** *A system of linear equations has a solution if and only if the rank of the coefficient matrix $A$ is equal to the rank of the extended matrix $(A|b)$.*

*Proof.* Assume that the system of linear equations has at least one solution. It means that $b$ is expressed as a linear combination of columns of $A$ and, thus, $\mathrm{rk}(A|b)$ must be the same as $\mathrm{rk}(A)$ since linear combination doesn't change the rank. Vice versa, assume that $\mathrm{rk}(A|b) = \mathrm{rk}(A) = k$. It means that the last equation in the echelon form of the extended matrix is $a_k x_k + a_{k+1} x_{k+1} + \cdots + a_n x_n = b_k$. We can set $x_{k+1} = \cdots = x_n = 0$ and use theorem 11 to find the solution for $x_1, \ldots, x_k$ as $\mathrm{rk}(A) = k$ and the corresponding minor is non-zero. □

**Corollary 12.1.** Consider the notation of theorem 11. If $\Delta = 0$ but $\Delta_i \neq 0$ for some $i$ then the system of linear equations has no solutions.

The connection between the number of solutions to a system of linear equations and the determinant of the matrix that system leads us to the following important definition

**Definition 13.** A square n×n matrix $A$ is called *regular* (or *non-degenerate*) if $\det(A) \neq 0$. Otherwise, it is called *singular* (or *degenerate*).

Out of these four terms, we will preferentially use the term non-degenerate when describing the properties of matrix determinant. The term itself stems from the property of the system of linear equations, which *degenerates* when some its equations become linearly dependent. If you pick a set of $n^2$ numbers at random and put them into a square table, in "most cases" (that is, with probability 1) it will give a non-degenerate matrix because a set of 3D planes in general position intersects by one point, but "sometimes" (with probability 0) it will become special, i.e. *singular*, and give many or no solutions. This explains the other pair of terms. It follows immediately from definition 13 and theorem 12 that

**Corollary 12.2.** A system of linear equations with $n$ variables and $n$ equations has a unique solution if and only if its matrix is non-degenerate.

**Corollary 12.3.** A n×n matrix $A$ is non-degenerate if and only if $A$ has the maximum possible rank, i.e., if $\mathrm{rk}(A) = n$.

**Corollary 12.4.** A n×n matrix $A$ is non-degenerate if and only if its rows (or columns) are linearly independent.

## 4.7 Exercises

PROBLEM 4.1. Evaluate the following determinants

(a) $\begin{vmatrix} -4 & -4 & 0 & 0 & 4 \\ 6 & -3 & -12 & 8 & 4 \\ 0 & -9 & -12 & 0 & 0 \\ 6 & 4 & 12 & 8 & 2 \\ 12 & 6 & 12 & 12 & 0 \end{vmatrix}$ (b) $\begin{vmatrix} 9 & -7 & -2 \\ -10 & -12 & 2 \\ 4 & -11 & -1 \end{vmatrix}$ (c) $\begin{vmatrix} 9 & 4 & 1 \\ -12 & -4 & 4 \\ -12 & 0 & 12 \end{vmatrix}$

(d) $\begin{vmatrix} 1 & 4 & 11 & -9 & 4 \\ -1 & -4 & -2 & 6 & 0 \\ -9 & -4 & -2 & -2 & 0 \\ -8 & 1 & 5 & 7 & -4 \\ 7 & -4 & -11 & 5 & 0 \end{vmatrix}$ (e) $\begin{vmatrix} 0 & 0 & -6 \\ 1 & -1 & 8 \\ 10 & -2 & 4 \end{vmatrix}$ (f) $\begin{vmatrix} -2 & 11 & 1 \\ 1 & 9 & -8 \\ 2 & 6 & -10 \end{vmatrix}$

(g) $\begin{vmatrix} 0 & 3 & 0 & -11 & -1 \\ 12 & -2 & -2 & 10 & 2 \\ -6 & -3 & 0 & 7 & -4 \\ -6 & -4 & 1 & 12 & 7 \\ 6 & -2 & -1 & 10 & -10 \end{vmatrix}$ (h) $\begin{vmatrix} 8 & -2 & 1 \\ -9 & 3 & 12 \\ 12 & -3 & 3 \end{vmatrix}$ (i) $\begin{vmatrix} -10 & 4 & 0 \\ 1 & -7 & -1 \\ -7 & -10 & -2 \end{vmatrix}$

(j) $\begin{vmatrix} 7 & 10 & -4 & -4 \\ 1 & -12 & 8 & 6 \\ -3 & 6 & 12 & 6 \\ -5 & -8 & 12 & 8 \end{vmatrix}$ (k) $\begin{vmatrix} 5 & 11 & 7 \\ -2 & 3 & -10 \\ -4 & -4 & -10 \end{vmatrix}$ (l) $\begin{vmatrix} 3 & -4 & 0 & -6 \\ 12 & 1 & 0 & -2 \\ -12 & -2 & 0 & 0 \\ -6 & 10 & 4 & -1 \end{vmatrix}$

PROBLEM 4.2. Evaluate ranks of the following matrices by any appropriate methods.

(a) $\begin{pmatrix} 12 & 0 & 7 & 5 & 4 \\ 24 & 0 & 14 & 10 & 8 \\ -8 & 16 & -2 & -6 & 0 \\ -12 & -6 & -8 & -4 & -5 \end{pmatrix}$ (b) $\begin{pmatrix} -20 & -10 & 13 & -1 & 13 \\ -12 & -9 & 12 & 0 & 9 \\ 12 & 4 & -5 & 1 & -7 \\ 20 & -10 & 15 & 5 & -5 \end{pmatrix}$

(c) $\begin{pmatrix} 16 & 18 & -6 & -5 & 8 \\ -16 & -18 & 6 & 8 & -14 \\ 19 & 12 & -8 & 3 & -8 \\ 5 & -9 & -17 & -14 & -3 \end{pmatrix}$ (d) $\begin{pmatrix} 3 & -2 & -4 & 2 & 0 \\ -10 & 2 & 8 & -4 & 2 \\ -19 & -6 & 4 & -2 & 8 \\ -17 & -12 & -4 & 2 & 10 \end{pmatrix}$

(e) $\begin{pmatrix} -15 & -11 & 9 & 0 & -16 \\ -8 & -10 & 6 & -2 & -8 \\ -22 & -12 & 12 & 2 & -24 \\ 5 & 14 & -6 & 5 & 4 \end{pmatrix}$ (f) $\begin{pmatrix} -12 & -8 & -16 & 5 & -11 \\ 6 & -2 & 11 & 2 & -2 \\ -10 & 6 & -18 & 8 & 10 \\ -3 & -11 & -1 & -4 & -17 \end{pmatrix}$

Problem 4.3. Solve the following SLE's by using the Cramer's rule, when possible.

(a) $\begin{cases} x_1 + 3x_2 + 2x_3 = 6 \\ -4x_1 + x_2 + 3x_3 = 7 \\ -5x_1 + x_2 - x_3 = -10 \end{cases}$ (b) $\begin{cases} -3x_1 + 4x_2 - 5x_3 = -2 \\ -4x_1 - 2x_2 - 5x_3 = 7 \\ 4x_1 + x_2 - 5x_3 = 25 \end{cases}$

(c) $\begin{cases} -x_1 + x_2 + 2x_3 = -5 \\ 4x_1 + 3x_2 - 5x_3 = 12 \\ -4x_1 - 5x_2 = 11 \end{cases}$ (d) $\begin{cases} -5x_1 - 3x_2 + 2x_3 - 3x_4 = -11 \\ 2x_1 + x_2 + 4x_3 + 2x_4 = -10 \\ 4x_1 - 4x_2 - x_3 + 4x_4 = -26 \\ -5x_1 + 4x_2 + 2x_3 + 2x_4 = -8 \end{cases}$

(e) $\begin{cases} 3x_2 - 5x_3 - 2x_4 = -24 \\ 3x_1 + x_2 + 3x_3 + 2x_4 = 1 \\ -5x_1 + x_2 - 3x_3 - 3x_4 = 7 \\ 4x_1 + x_2 + 2x_3 - 5x_4 = -22 \end{cases}$ (f) $\begin{cases} -4x_1 - 4x_2 - 5x_3 = -6 \\ -3x_1 + 2x_2 - 5x_3 = 13 \\ -x_1 - 3x_2 + x_3 = -12 \end{cases}$

(g) $\begin{cases} -4x_1 - 2x_3 = -4 \\ -3x_1 - 2x_2 - 3x_3 = 1 \\ -x_1 - 2x_2 + x_3 = 3 \end{cases}$ (h) $\begin{cases} 3x_1 - x_2 - 5x_3 + 4x_4 = -2 \\ -x_1 - 3x_2 - 3x_3 = 4 \\ x_2 - 4x_3 - 5x_4 = -11 \\ 2x_1 + x_2 - x_3 + 3x_4 = -2 \end{cases}$

# 5 Matrix algebra

## 5.1 Matrix operations

The set of n×m matrices $M_{n,m}(R)$ is naturally equipped with the structure of a linear vector space with respect to the operation of matrix addition that is defined below. Let $A, B \in M_{n,m}(R)$,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{pmatrix}.$$

The matrices $A+B$ and $\lambda A$ are defined by component-wise addition of elements of $A$ and $B$ and component-wise multiplication of elements of $A$ by $\lambda$, respectively, i.e.,

$$A+B=\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2m}+b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \dots & a_{nm}+b_{nm} \end{pmatrix},$$

$$\lambda A=\begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1m} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{n1} & \lambda a_{n2} & \dots & \lambda a_{nm} \end{pmatrix}$$

**Example 5.1.**

$$A=\begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{pmatrix}, \qquad B=\begin{pmatrix} 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad A+B=\begin{pmatrix} 2 & 2 & 2 \\ 3 & 4 & 6 \end{pmatrix}.$$

The following properties of matrix addition and scaling are evident.

1. $\forall A,B,C \in M_{n,m}(\mathbb{R})\ \ (A+B)+C=A+(B+C)$
2. $\exists 0 \in M_{n,m}(\mathbb{R}) : \forall A \in M_{n,m}(\mathbb{R})\ \ A+0=0+A=A$
3. $\forall A \in M_{n,m}(\mathbb{R})\ \exists (-A) \in M_{n,m}(\mathbb{R}) :\ \ A+(-A)=(-A)+A=0$
4. $A+B=B+A$ for $\forall A,B \in M_{n,m}(\mathbb{R})$
5. $\forall \lambda \in \mathbb{R}\ \forall A,B \in M_{n,m}(\mathbb{R})\ \lambda(A+B)=\lambda A+\lambda B$
6. $\forall \lambda,\mu \in \mathbb{R}\ \forall A \in M_{n,m}(\mathbb{R})\ (\lambda\mu)A=\lambda(\mu)A$

Thus, $M_{n,m}(\mathbb{R})$ is a linear vector space[5] with $\dim M_{n,m}(\mathbb{R}) = nm$. In contrast to matrix addition, matrix multiplication is generally defined for matrices of different sizes.

Let $A \in M_{n,k}(\mathbb{R})$ and $B \in M_{k,m}(\mathbb{R})$. Note that the number of columns of the first matrix is equal to the number of rows of the second matrix. Such matrices are called *compatible*. Then, the matrix $C = A \times B = (c_{ij})$ is defined by the formula

$$c_{ij} = \sum_{s=1}^{k} a_{is} \times b_{sj}.$$

**Example 5.2.**

$$A=\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad B=\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

[5] Can you find a natural base in this space?

$$c_{11} = \sum_{s=1}^{2} a_{1s}b_{s1} = a_{11}b_{11} + a_{12}b_{21} = 1 \times 1 + 2 \times 4 = 9,$$

$$c_{12} = \sum_{s=1}^{2} a_{1s}b_{s2} = a_{11}b_{12} + a_{12}b_{22} = 1 \times 2 + 2 \times 5 = 12,$$

$$\dots$$

$$c_{23} = \sum_{s=1}^{2} a_{2s}b_{s3} = a_{21}b_{13} + a_{22}b_{23} = 3 \times 3 + 4 \times 6 = 33,$$

$$A \times B = \begin{pmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \end{pmatrix}.$$

Matrix multiplication is a binary operation which can be though of as a mapping $f : M_{n,k}(\mathbb{R}) \times M_{k,m}(\mathbb{R}) \to M_{n,m}(\mathbb{R})$ with the following properties.

1. $\forall A \in M_{n,k}(\mathbb{R}) \, \forall B \in M_{k,l}(\mathbb{R}) \, \forall C \in M_{l,m}(\mathbb{R}) \;\; (A \times B) \times C = A \times (B \times C)$
2. $\forall A \in M_{n,k}(\mathbb{R}) \, \forall B, C \in M_{k,m}(\mathbb{R}) \;\; A \times (B + C) = A \times B + A \times C$
3. $\forall \lambda \in \mathbb{R} \, \forall A \in M_{n,k}(\mathbb{R}) \, \forall B \in M_{k,m}(\mathbb{R}) \;\; (\lambda A) \times B = \lambda(A \times B)$

Note that, in general, $A \times B \neq B \times A$, i.e., matrix multiplication is *not commutative.* This is the most striking difference of matrix multiplication from the multiplication of numbers. So, from now on we will pay special attention to the order of terms in matrix products. For instance, we distinguish between left and right multiplication, where $A$ multiplied by $B$ *from the right* means $A \times B$, while $A$ multiplied by $B$ *from the left* means $B \times A$.

**Example 5.3.**

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad A \times B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad B \times A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

**Lemma 13.** $(A \times B)^T = B^T \times A^T$

*Proof.* Consider $A \times B = C = (c_{ij})$. That is, $c_{ij} = \sum_k a_{ik}b_{kj}$. Then, $C^T = (c_{ji})$ and $c_{ji} = \sum_k a_{jk}b_{ki} = \sum_k b_{ki}a_{jk} = \sum_k \tilde{b}_{ik}\tilde{a}_{kj}$, where $\tilde{b}_{ij}$ and $\tilde{a}_{ij}$ are the elements of $B^T$ and $A^T$, respectively. Therefore, $C^T = B^T \times A^T$. □

## 5.2 Inverse matrix

In this section we consider only square matrices. Let $A \in M_{nn}(\mathbb{R})$ be such a matrix. The notion of *inverse matrix* is very similar to the notion of inverse (reciprocal) number. The reciprocal to $x$ is $x^{-1}$, which is defined as a number such that $x \times x^{-1} = 1$. For matrices, the role of 1 is played by the identity matrix

$$E_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in M_{nn}(\mathbb{R}),$$

which is the neutral element of matrix multiplication, i.e., $A \times E = E \times A = A$ for every $A \in M_{nn}(\mathbb{R})$.

**Definition 14.** Let $A$ be a square n×n matrix. The matrix $A^{-1} \in M_{nn}(\mathbb{R})$ is called the *inverse matrix* for $A$ if $A \times A^{-1} = A^{-1} \times A = E$, where $E$ is the identity matrix.

**Lemma 14.** $(A \times B)^{-1} = B^{-1} \times A^{-1}$.

*Proof.* $(B^{-1} \times A^{-1}) \times (A \times B) = B^{-1} \times (A^{-1} \times A) \times B = B^{-1} \times B = E$. Similarly, $(A \times B) \times (B^{-1} \times A^{-1}) = A \times (B \times B^{-1}) \times A^{-1} = A \times A^{-1} = E$. □

**Lemma 15.** $(A^T)^{-1} = (A^{-1})^T$.

*Proof.* According to lemma 13, $(A^{-1})^T \times A^T = (A \times A^{-1})^T = E$ and $A^T \times (A^{-1})^T = (A^{-1} \times A)^T = E$. Thus, $(A^{-1})^T$ is the inverse to $A^T$. □

## 5.3 Inverse matrix by Gauss elimination

The following theorem gives the necessary and sufficient condition of existence of the inverse matrix.

**Theorem 16.** $A^{-1}$ *exists if and only if* $\det(A) \neq 0$

*Proof.* The necessary part follows from the definition: $\det(A \times A^{-1}) = \det(A)\det(A^{-1}) = 1$ and, therefore, $\det(A)$ cannot be equal to zero.

Now we prove the sufficient condition. Assume that

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix},$$

and $C = A \times B$. We need $C = E$. Thus, for the first column of $C$ we get

$$c_{11} = a_{11}b_{11} + a_{12}b_{21} + \dots + a_{1n}b_{n1} = 1,$$

$$c_{21} = a_{21}b_{11} + a_{22}b_{21} + \dots + a_{2n}b_{n1} = 0,$$

$$c_{n1} = a_{n1}b_{11} + a_{n2}b_{21} + \dots + a_{nn}b_{n1} = 0,$$

That is, in order to solve for $b_{11}, \dots, b_{n1}$, we have to solve the SLE

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & 1 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 \end{array}\right)$$

Similarly, the second column of $C$ gives

$$c_{12} = a_{11}b_{12} + a_{12}b_{22} + \dots + a_{1n}b_{n2} = 0,$$

$$c_{22} = a_{21}b_{12} + a_{22}b_{22} + \dots + a_{2n}b_{n2} = 1,$$

$$c_{n2} = a_{n1}b_{12} + a_{n2}b_{22} + \dots + a_{nn}b_{n2} = 0,$$

i.e., in order to get $b_{12}, \dots, b_{n2}$, we need to solve for

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 \end{array}\right).$$

A similar SLE can be written for $b_{1n}, \dots, b_{nn}$. Note that all these SLEs have the same matrix and differ only by the column of free terms. Thus, we can solve them all simultaneously because the sequence of operations that are applied to the matrix of SLE will be the same for each column in the right-hand side. In other words, we can write

$$\left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array}\right),$$

or, shortly, $(A|E)$ and apply Gauss elimination procedure to transform $A$ to $E$. Then, the matrix on the right of the vertical bar will be $A^{-1}$. According to theorem 11, if $\det(A) \neq 0$ then the solution $(A|E)$ exists and is unique. □

*Remark.* The proof of this theorem can be used to calculate $A^{-1}$ numerically.

**Example 5.4.** Given $A = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$, compute $A^{-1}$

**Solution.**

$$\left(\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 2 & -1 & 1 \\ 2 & 3 & 1 & 0 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 2 & -1 & 1 \\ 0 & -1 & 3 & -2 \end{array}\right) \to$$

$$\left(\begin{array}{cc|cc} 1 & 2 & -1 & 1 \\ 0 & 1 & -3 & 2 \end{array}\right) \to \left(\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 0 & 5 & -3 \\ 0 & 1 & -3 & 2 \end{array}\right)$$

Verify by multiplication: $\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \times \begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

That is, $A^{-1} = \begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix}$.

**Example 5.5.** Given $A = \begin{pmatrix} 2 & 1 & -1 \\ -1 & 1 & 2 \\ 0 & 3 & 1 \end{pmatrix}$, find $A^{-1}$.

**Solution.**

$$\left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ -1 & 1 & 2 & 0 & 1 & 0 \\ 0 & 3 & 1 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & -1 & -2 & 0 & -1 & 0 \\ 0 & 3 & 3 & 1 & 2 & 0 \\ 0 & 3 & 1 & 0 & 0 & 1 \end{array}\right) \to$$

$$\left(\begin{array}{ccc|ccc} 1 & -1 & -2 & 0 & -1 & 0 \\ 0 & 1 & 1 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & 1 & \frac{1}{2} & 1 & -\frac{1}{2} \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 1 & -1 \\ 0 & 1 & 0 & -\frac{1}{6} & -\frac{1}{3} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & 1 & -\frac{1}{2} \end{array}\right) \to$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{5}{6} & \frac{2}{3} & -\frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{6} & -\frac{1}{3} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & 1 & -\frac{1}{2} \end{array}\right).$$

Check:

$$\begin{pmatrix} 2 & 1 & -1 \\ -1 & 1 & 2 \\ 0 & 3 & 1 \end{pmatrix} \times \begin{pmatrix} \frac{5}{6} & \frac{2}{3} & -\frac{1}{2} \\ -\frac{1}{6} & -\frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 1 & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}..$$

## 5.4 Inverse matrix by algebraic complements

In this section we discuss the algorithm for calculation of $A^{-1}$ which is based on the use of determinants. The inverse matrix can be calculated in five steps.

1. Calculate $\det(A)$. Proceed to the next step unless $\det(A) = 0$.
2. For each $i$ and $j$, compute the value of algebraic complement $M_{ij}$ (see definition in section 4.4 on page 29) and collect them in a new matrix.
3. Transpose the matrix obtained in step 2.
4. Multiply each element of the transpose matrix by $(-1)^{i+j}$, i.e., change signs according to the chess-board pattern:
$$\begin{pmatrix} + & - & + & - & + \\ - & + & - & + & - \\ + & - & + & - & + \\ - & + & - & + & - \end{pmatrix}$$
5. Divide each element of the obtained matrix by $\det(A)$

**Example 5.6.** Find the inverse matrix for $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 3 \end{pmatrix}$

**Solution.**

1. det A= 3+0+0-1-12-0=-10
2. Compute algebraic complements (also called *cofactor matrix*)

$$M_{11} = \begin{vmatrix} 1 & 0 \\ 0 & 3 \end{vmatrix} = 3, \quad M_{12} = 6, \quad M_{13} = -1$$

$$M_{21} = \begin{vmatrix} 2 & 1 \\ 0 & 3 \end{vmatrix} = 6, \quad M_{22} = 2, \quad M_{23} = -2$$

$$M_{31} = \begin{vmatrix} 2 & 1 \\ 1 & 0 \end{vmatrix} = -1, \quad M_{32} = -2, \quad M_{33} = -3$$

3. Transpose

$$\begin{pmatrix} 3 & 6 & -1 \\ 6 & 2 & -2 \\ -1 & -2 & -3 \end{pmatrix}^T = \begin{pmatrix} 3 & 6 & -1 \\ 6 & 2 & -2 \\ -1 & -2 & -3 \end{pmatrix}.$$

4. Change signs

$$\begin{pmatrix} 3 & -6 & -1 \\ -6 & 2 & 2 \\ -1 & 2 & -3 \end{pmatrix}.$$

5. Divide each element by $\det(A)$

$$A^{-1} = \begin{pmatrix} -0.3 & 0.6 & 0.1 \\ 0.6 & -0.2 & -0.2 \\ 0.1 & -0.2 & 0.3 \end{pmatrix}..$$

6. Check:

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 3 \end{pmatrix} \times \begin{pmatrix} -0.3 & 0.6 & 0.1 \\ 0.6 & -0.2 & -0.2 \\ 0.1 & -0.2 & 0.3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

## 5.5 Matrix equations

A matrix equation is written similarly to a numeric equation, like $AX = B$, with the exception that $A$ and $B$ are matrices, and the unknown $X$ is a matrix, too (for simplicity of notation, we omit the multiplication sign, $\times$, as we often do for multiplication of numbers). When you solve numeric equations, for instance $2x = 3$, you simply divide the right-hand side by the factor in the left-hand side, i.e., $x = 3/2$. Division is not that easy for matrices, because it is equivalent to multiplication by the inverse matrix, which can be done from the right or from the left, each with different outcomes.

In order to "cancel out" the matrix $A$ in the left-hand side, one needs to multiply the matrix equation $AX = B$ by $A^{-1}$ from the left, not from the right. Then $A^{-1}AX = EX = X = A^{-1}B$. Similarly, the solution to the equation $XA = B$ is $X = XAA^{-1} = BA^{-1}$.

**Example 5.7.** Solve the matrix equation

$$\begin{pmatrix} 5 & -1 & 3 \\ 5 & 0 & 4 \\ 6 & -1 & 4 \end{pmatrix} X = \begin{pmatrix} -2 & 4 \\ 4 & 6 \\ 6 & 0 \end{pmatrix}.$$

**Solution.** First, we calculate the inverse matrix

$$A^{-1} = \begin{pmatrix} 4 & 1 & -4 \\ 4 & 2 & -5 \\ -5 & -1 & 5 \end{pmatrix}.$$

Then,

$$X = A^{-1}B = \begin{pmatrix} 4 & 1 & -4 \\ 4 & 2 & -5 \\ -5 & -1 & 5 \end{pmatrix} \times \begin{pmatrix} -2 & 4 \\ 4 & 6 \\ 6 & 0 \end{pmatrix} = \begin{pmatrix} -28 & 22 \\ -30 & 28 \\ 36 & -26 \end{pmatrix}.$$

One can check by direct calculation that

$$\begin{pmatrix} 5 & -1 & 3 \\ 5 & 0 & 4 \\ 6 & -1 & 4 \end{pmatrix} \times \begin{pmatrix} -28 & 22 \\ -30 & 28 \\ 36 & -26 \end{pmatrix} = \begin{pmatrix} -2 & 4 \\ 4 & 6 \\ 6 & 0 \end{pmatrix}.$$

□

There is another way to solve matrix equations, one that is based on the direct calculations. Consider, for instance, the matrix equation $AX = B$. Similarly to what we have been doing in section 5.3, the matrix product of $A$ with the first column of $X$ gives the following linear equations

$$\begin{cases} a_{11}x_{11} + a_{12}x_{21} + \ldots + a_{1n}x_{n1} = b_{11} \\ a_{21}x_{11} + a_{22}x_{21} + \ldots + a_{2n}x_{n1} = b_{21} \\ a_{n1}x_{11} + a_{n2}x_{21} + \ldots + a_{nn}x_{n1} = b_{n1} \end{cases}.$$

Thus, in order to solve for the first column of $X$, we need to solve SLE with the matrix

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_{11} \\ a_{21} & a_{22} & \ldots & a_{2n} & b_{21} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} & b_{n1} \end{array}\right).$$

Similarly, to solve for the second column of $X$, we need to solve the equations

$$\begin{cases} a_{11}x_{12} & + & a_{12}x_{22} & + & \dots & + & a_{1n}x_{n2} & = b_{12} \\ a_{21}x_{12} & + & a_{22}x_{22} & + & \dots & + & a_{2n}x_{n2} & = b_{22} \\ a_{n1}x_{12} & + & a_{n2}x_{22} & + & \dots & + & a_{nn}x_{n2} & = b_{n2} \end{cases},$$

which yields SLE matrix, which differs from the previous one only by the column of free terms

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_{12} \\ a_{21} & a_{22} & \dots & a_{2n} & b_{22} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_{n2} \end{array}\right).$$

We can solve all these systems simultaneously by writing all $b$-terms to the right of the vertical bar a(s in section 5.3) and applying Gauss elimination

$$\left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & b_{11} & b_{12} & \vdots & b_{1m} \\ a_{21} & a_{22} & \dots & a_{2n} & b_{21} & b_{22} & \vdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_{n1} & b_{n2} & \vdots & b_{nm} \end{array}\right).$$

**Example 5.8.** Solve the matrix equation from the example 5.7 by Gauss elimination.

**Solution.**

$$\left(\begin{array}{ccc|cc} 5 & -1 & 3 & -2 & 4 \\ 5 & 0 & 4 & 4 & 6 \\ 6 & -1 & 4 & 6 & 0 \end{array}\right) \xrightarrow{[3]\mapsto[3]-[1]} \left(\begin{array}{ccc|cc} 1 & 0 & 1 & 8 & -4 \\ 5 & -1 & 3 & -2 & 4 \\ 5 & 0 & 4 & 4 & 6 \end{array}\right) \xrightarrow{[2]\mapsto[2]-[1]\times 5}$$

$$\left(\begin{array}{ccc|cc} 1 & 0 & 1 & 8 & -4 \\ 0 & -1 & -2 & -42 & 24 \\ 5 & 0 & 4 & 4 & 6 \end{array}\right) \xrightarrow{[3]\mapsto[3]-[1]\times 5} \left(\begin{array}{ccc|cc} 1 & 0 & 1 & 8 & -4 \\ 0 & 1 & 2 & 42 & -24 \\ 0 & 0 & 1 & 36 & -26 \end{array}\right) \xrightarrow{[1]\mapsto[1]-[3]}$$

$$\left(\begin{array}{ccc|cc} 1 & 0 & 0 & -28 & 22 \\ 0 & 1 & 2 & 42 & -24 \\ 0 & 0 & 1 & 36 & -26 \end{array}\right) \xrightarrow{[2]\mapsto[2]-[3]\times 2} \left(\begin{array}{ccc|cc} 1 & 0 & 0 & -28 & 22 \\ 0 & 1 & 0 & -30 & 28 \\ 0 & 0 & 1 & 36 & -26 \end{array}\right).$$

□

## 5.6 Elementary transformations matrices

All of the transformations discussed in section 1.2 can be expressed in terms of matrix multiplication with the aid of special matrices called *elementary transformation matrices.*

The Type I transformation, in which the $j^{\text{th}}$ row is multiplied by some constant $\lambda$ and added to the $i^{\text{th}}$ row is equivalent to the multiplication (from the left) of the original matrix by the matrix

$$U_{ij}(\lambda) = \begin{pmatrix} 1 & 0 & \lambda & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix},$$

where $\lambda$ is located in the intersection of the $i^{\text{th}}$ row and $j^{\text{th}}$ column, for instance in the case of $3\times3$ matrices,

$$U_{13}(\lambda) = \begin{pmatrix} 1 & 0 & \lambda \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The Type II transformation, in which the $i^{\text{th}}$ row is multiplied by a non-zero constant $\lambda$, is equivalent to the multiplication (again, from the left) of the original matrix by

$$P_i(\lambda) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix},$$

where $\lambda$ is located in the $i^{\text{th}}$ row on the diagonal, for instance,

$$P_2(\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The Type III transformation, in which the $i^{\text{th}}$ row and the $j^{\text{th}}$ row of the original matrix are interchanged, can be expressed as multiplication (again, from the left) by

$$T_{ij}(\lambda) = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 \end{pmatrix},$$

where the interchanged 1's are located in the intersection of the $i^{\text{th}}$ and $j^{\text{th}}$ row and $i^{\text{th}}$ and $j^{\text{th}}$ column, for instance

$$T_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

**Example 5.9.** Let A be a 3×3 matrix of the general form

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

then $U_{13}(\lambda) \times A =$

$$= \begin{pmatrix} 1 & 0 & \lambda \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} + \lambda a_{31} & a_{12} + \lambda a_{32} & a_{13} + \lambda a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

$$P_2(\lambda) \times A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ \lambda a_{21} & \lambda a_{22} & \lambda a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \text{ and}$$

$$T_{12} \times A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Similarly, one can show that the multiplication from the right by $U_{ji}(\lambda)$, $P_i(\lambda)$, and $T_{ij}$ leads to Type I, Type II, and Type III transformations of columns, respectively[6]. Note that the multiplication by $U_{ij}(\lambda)$ doesn't change $\det(A)$ because it is a type I transformation, while the multiplication by $T_{ij}$ leads to the change of sign of $\det(A)$. Then, we will sometimes use the matrix

$$R_{ij} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & -1 & \dots & 0 \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 \end{pmatrix},$$

where the 1 and -1 are located in the intersection of the $i^{\text{th}}$ and $j^{\text{th}}$ row and $i^{\text{th}}$ and $j^{\text{th}}$ column. Its important feature is that the multiplication of a matrix by $R_{ij}$ from the left or right doesn't change its determinant. We will use the elementary transformation matrices to prove the following theorem.

**Theorem 17.**

$$\det(A \times B) = \det(A)\det(B)$$

*Proof.* First, the proof is trivial when $A$ or $B$ is a degenerate matrix. Indeed, if $A$ or $B$ is degenerate then $A \times B$ is degenerate, too, because its rows are linear combinations of a linearly dependent set of vectors and we get both $\det(A \times B)$ and $\det(A)\det(B)$ equal to zero.

[6]Spend some time with paper and pencil to check that it is, indeed, true!

Next, we prove this statement for the particular case when both $A$ and $B$ are upper-triangular matrices. Assume that

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} & \dots & b_{1n} \\ 0 & b_{22} & b_{23} & \dots & b_{2n} \\ 0 & 0 & b_{33} & \dots & b_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & b_{nn} \end{pmatrix}.$$

Then, by definition of the matrix multiplication, we get

$$\begin{pmatrix} a_{11}b_{11} & * & * & \dots & * \\ 0 & a_{22}b_{22} & * & \dots & * \\ 0 & 0 & a_{33}b_{33} & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn}b_{nn} \end{pmatrix},$$

and, since the determinant of an upper-triangular matrix is the product of its diagonal entries (theorem 8), we get $\det(A \times B) = a_{11}b_{11}a_{22}b_{22}\dots a_{nn}b_{nn} = a_{11}a_{22}\dots a_{nn}b_{11}b_{22}\dots b_{nn} = \det(A)\det(B)$.

The general case can be reduced to the case of upper-triangular matrices by elementary transformations of rows, for the matrix $A$, and by elementary transformations of columns, for the matrix $B$. Indeed, the matrix $A$ can be represented as a product $A = A_1 \times A_2 \times \dots \times A_k \times X$, where $A_i$ is one of the matrices $U_{ij}(\lambda)$ or $R_{ij}$ which don't affect the determinant, while $X$ is an upper-triangular matrix such that $\det(A) = \det(X)$. Similarly, we can transform $B$ by elementary transformations of columns to the form $Y \times B_l \times \dots \times B_2 \times B_1$, where $B_j$ are matrices $U_{ij}(\lambda)$ or $R_{ij}$ and $Y$ is also an upper-triangular matrix such that $\det(B) = \det(Y)$. Then, $\det(A \times B) = \det(A_1 \times A_2 \times \dots \times A_k \times X \times Y \times B_l \times \dots \times B_2 \times B_1) = \det(X \times Y) = \det(X)\det(Y) = \det(A)\det(B)$. □

## 5.7 Exercises

Problem 5.1. Find the inverse matrix for each of the following matrices

(a) $\begin{pmatrix} -1 & -2 & 4 & 11 & -2 \\ -4 & 0 & 9 & -7 & -3 \\ 0 & 1 & -1 & -8 & 0 \\ -5 & -1 & 12 & -3 & -5 \\ 4 & 0 & -9 & 9 & 4 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 0 & -7 & -1 \\ 0 & 0 & 1 & 0 \\ -5 & 3 & -9 & 2 \\ -12 & 5 & -3 & 5 \end{pmatrix}$

(c) $\begin{pmatrix} -6 & 7 & 7 & -6 \\ -1 & 2 & 0 & 0 \\ -6 & 7 & 8 & -6 \\ 2 & -3 & -1 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 7 & 4 & 6 & 3 \\ 4 & -2 & -2 & 1 \\ 12 & 6 & 9 & 5 \\ 10 & 3 & 5 & 4 \end{pmatrix}$

(e) $\begin{pmatrix} 3 & -2 & -4 \\ -2 & 1 & 3 \\ -6 & 3 & 8 \end{pmatrix}$ (f) $\begin{pmatrix} 11 & -11 & 10 & 10 \\ 0 & -1 & 1 & 1 \\ -3 & 5 & -4 & -5 \\ 7 & -4 & 4 & 3 \end{pmatrix}$

(g) $\begin{pmatrix} -11 & -10 & -6 \\ -6 & -5 & -3 \\ -10 & -8 & -5 \end{pmatrix}$ (h) $\begin{pmatrix} -3 & 10 & -1 & 0 & -1 \\ 4 & -12 & 1 & 0 & 1 \\ 1 & -2 & 3 & 0 & -2 \\ -1 & 0 & 5 & 1 & -2 \\ 0 & -3 & 7 & 1 & -3 \end{pmatrix}$

(i) $\begin{pmatrix} 3 & -7 & -11 & -7 & 5 \\ 7 & -1 & 11 & -8 & 3 \\ -5 & 2 & -5 & 6 & -3 \\ 5 & 1 & 12 & -5 & 1 \\ 1 & 4 & 12 & 1 & -2 \end{pmatrix}$ (j) $\begin{pmatrix} 1 & 2 & -1 & -1 & -2 \\ -1 & -8 & 10 & 0 & 5 \\ 1 & 1 & -2 & -4 & -2 \\ -1 & 0 & 2 & 6 & 2 \\ 1 & 2 & -5 & -6 & -3 \end{pmatrix}$

Problem 5.2. Solve the following matrix equations

(a) $\begin{pmatrix} -3 & 3 & 5 \\ 1 & -1 & -2 \\ -7 & 6 & 10 \end{pmatrix} X = \begin{pmatrix} -1 & -5 & -4 \\ 3 & 2 & -2 \\ -1 & 2 & 2 \end{pmatrix}$

(b) $\begin{pmatrix} 6 & 4 & 1 & -6 \\ -5 & -4 & -1 & 6 \\ 8 & 3 & 1 & -4 \\ 10 & 0 & 1 & 1 \end{pmatrix} X = \begin{pmatrix} 1 & -1 & 0 & -4 \\ -4 & -5 & -4 & -2 \\ -2 & -2 & 2 & 1 \\ -1 & 2 & 1 & 3 \end{pmatrix}$

(c) $X \begin{pmatrix} 0 & -3 & 11 & 0 & 9 \\ 1 & 0 & -5 & -1 & -3 \\ 1 & 0 & 5 & 2 & 5 \\ -1 & 4 & -2 & 3 & -3 \\ -1 & -5 & 2 & -5 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 3 & -1 & 3 \\ 4 & 1 & -3 & 2 & 3 \\ -4 & 1 & -4 & -1 & -4 \\ -2 & 4 & 1 & 0 & 3 \end{pmatrix}$

(d) $X \begin{pmatrix} 7 & -6 & -8 \\ -1 & 1 & 1 \\ 3 & -2 & -3 \end{pmatrix} = \begin{pmatrix} 1 & -2 & -3 \\ 1 & 2 & -3 \\ -4 & 3 & 2 \\ 4 & 1 & -3 \\ 1 & -5 & -5 \end{pmatrix}$

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

(f) $\begin{pmatrix} 8 & -12 & -1 & 0 \\ -3 & 7 & 2 & -1 \\ -7 & 6 & -2 & 2 \\ 0 & 5 & 3 & -2 \end{pmatrix} X = \begin{pmatrix} 2 & 3 & 2 & -2 & -4 \\ 4 & -3 & 1 & 1 & 4 \\ -5 & -3 & 2 & 2 & 4 \\ 3 & 4 & -2 & -5 & -2 \end{pmatrix}$

# 6 Linear operators

The notion of a linear operator is very general and has many applications in all parts of mathematics, including calculus, statistics, and optimization theory. Commonly used synonyms of a linear operator are linear map, linear mapping, linear transformation, or linear function.

**Definition 15.** Let $V$ and $W$ be linear vector spaces. The mapping $f : V \to W$ is said to be a *linear operator* if

1. $\forall x, y \in V \;\; f(x + y) = f(x) + f(y)$
2. $\forall x \in V, \forall \lambda \in \mathbb{R} \;\; f(\lambda x) = \lambda f(x)$,

i.e., $f$ preserves the operations of vector addition and scalar multiplication. Below we discuss several examples of mapping which are (or are not) linear operators.

1. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = cx$ is a linear operator as $c(x + y) = cx + cy$ and $c(\lambda x) = (c\lambda)x$.
2. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ is NOT a linear operator because $(x + y)^2 \neq x^2 + y^2$.
3. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x+1$ is NOT a linear operator because $(x+1)+(y+1) \neq (x+y)+1$.
4. The linear operator of differentiation by $x$, i.e., $\frac{\partial}{\partial x} : C^\infty[a, b] \to C^\infty[a, b]$ such that $\frac{\partial}{\partial x}(f) = f'(x)$, which is defined on the set of infinitely smooth functions (page 15) is a linear operator.
5. The linear operator of integration over a segment $[a, b]$, i.e., $I : C[a, b] \to \mathbb{R}$ such that $I(f) = \int_a^b f(x)dx$ is a linear operator as $\int_a^b (f + g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx$ and $\int_a^b (\lambda f)(x)dx = \lambda \int_a^b f(x)dx$.
6. The linear operator of expected value acts on the linear space of random numbers to $\mathbb{R}$ so that $E(X + Y) = E(X) + E(Y)$ and $E(\lambda X) = \lambda E(X)$.
7. $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (y, x)$ is a linear operator of symmetry with respect to the bisector line of the first quadrant, as the vector symmetric to a sum of two vectors is equal to the sum of their symmetric images (the same is true for scaling).
8. $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (y, -x)$ is a linear operator of rotation 90° counterclockwise (check it!), as the rotated sum of two vectors is equal to the sum of rotations.
9. $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x$ is a linear operator of projection onto the $x$-axis, because the projection of the sum of two vectors is equal to the sum of their projections.

10. $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \sqrt{x^2 + y^2}$ is NOT a linear operator, because the length of the sum of two vectors is not equal to the sum of their lengths.

## 6.1 Matrix of a linear operator

We now explore the connection between linear operators and matrices. It turns out, that each linear operator can be described with a matrix, but the problem is that this matrix is basis-dependent.

Consider a vector space $V$ with a base $e_1, e_2, \ldots, e_n$. Since each vector of $V$ can be expressed as a linear combination of $e_1, e_2, \ldots, e_n$, we get $x = x_1e_1 + x_2e_2 + \cdots + x_ne_n$, where $x_1, x_2, \ldots, x_n$ are the coordinates of $x$ in the base $e_1, e_2, \ldots, e_n$. If $A : V \to W$ is a linear operator then $A(x) = A(x_1e_1+x_2e_2+\cdots+x_ne_n) = x_1A(e_1)+x_2A(e_2)+\cdots+x_nA(e_n)$. Here, $A(e_1), A(e_2), \ldots, A(e_n)$ are the images of base vectors under the map $f$. Note that they are the same for all vectors $x$. Chose the base $f_1, f_2, \ldots, f_m$ in the vector space $W$ and decompose vectors $A(e_1), A(e_2), \ldots, A(e_n)$ by that base[7].

$$\begin{array}{ccccccccc} A(e_1) & = & a_{11}f_1 & + & a_{21}f_2 & + & \ldots & + & a_{m1}f_m, \\ A(e_2) & = & a_{12}f_1 & + & a_{22}f_2 & + & \ldots & + & a_{m2}f_m, \\ A(e_n) & = & a_{1n}f_1 & + & a_{2n}f_2 & + & \ldots & + & a_{mn}f_m. \end{array}$$

We have $A(x) = x_1A(e_1)+x_2A(e_2)+\cdots+x_nA(e_n) = x_1(a_{11}f_1+a_{21}f_2+\cdots+a_{m1}f_m)+x_2(a_{12}f_1+a_{22}f_2+\cdots+a_{m2}f_m)+\cdots+x_n(a_{1n}f_1+a_{2n}f_2+\cdots+a_{mn}f_m) = (a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n)f_1+(a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n)f_2+\ldots(a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n)f_m$. Thus, we can write the action of $y = A(x)$ by using matrix multiplication as

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{1mn} \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

where $x_1, x_2, \ldots, x_n$ are the coordinates of $x$ in the base $e_1, e_2, \ldots, e_n$ and $y_1, y_2, \ldots, y_m$ are the coordinates of $A(x)$ in the base $f_1, f_2, \ldots, f_m$. Note that it is the same base which gave rise to the matrix $(a_{ij})$ of the coordinates of $f(e_i)$. The matrix formed by $(a_{ij})$ will be called the *matrix of the linear operator* $A$ and will be denoted by the same capital letter as the linear operator. Its matrix is specific to the bases $e_1, e_2, \ldots, e_n$ and $f_1, f_2, \ldots, f_m$. In order to construct the matrix of a linear operator, one needs to take images of all base vectors of the space $V$, find their coordinates in the base of the space $W$, and write down the coordinates *as columns*. This gives the matrix of a linear operator.

**Example 6.1.** Let $A : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear operator of rotation $90^\circ$ counterclockwise on the plane. Consider $e_1 = f_1 = (1, 0)$ and $e_2 = f_2 = (0, 1)$. Then, $A(e_1) = f_2 = 0 \cdot f_1 + 1 \cdot f_2$ and $A(e_2) = -f_1 = (-1) \cdot f_1 + 0 \cdot f_2$. Therefore, the matrix of $A$ in these bases is

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

[7]Please pay attention to the notation of indices, which changed to make a more convenient expression below.

**Example 6.2.** Let $V$ be the vector space of polynomials of degree 2 or less and let $A : V \to V$ be the operator of differentiation by $x$. Consider the bases $e_1, e_2, e_3$ and $f_1, f_2, f_3$ such that $e_1 = f_1 = 1$, $e_2 = f_2 = x$, and $e_3 = f_3 = x^2$. Then, $A(e_1) = 0 = 0 \cdot f_1 + 0 \cdot f_2 + 0 \cdot f_3$, $A(e_2) = f_1 = 1 \cdot f_1 + 0 \cdot f_2 + 0 \cdot f_3$, and $A(e_3) = 2f_2 = 0 \cdot f_1 + 2 \cdot f_2 + 0 \cdot f_3$. Hence, the matrix of $A$ is

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

**Example 6.3.** Let $A : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear operator of rotation in 3D which induces a cyclic permutation of the axes $x$, $y$, and $z$. In other words, it is a rotation around the line $x = y = z$ by 120° clockwise when looking at the origin from the first octant[8]. Consider $e_1 = f_1 = (1, 0, 0)$, $e_2 = f_2 = (0, 1, 0)$, and $e_3 = f_3 = (0, 0, 1)$. Then, $A(e_1) = f_2 = 0 \cdot f_1 + 1 \cdot f_2 + 0 \cdot f_3$, $A(e_2) = f_3 = 0 \cdot f_1 + 0 \cdot f_2 + 1 \cdot f_3$, and $A(e_3) = f_1 = 1 \cdot f_1 + 0 \cdot f_2 + 0 \cdot f_3$. Therefore, the matrix of $A$ in these bases is

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

## 6.2 Change of base

In the previous section we considered two different vector spaces, $V$ and $W$, each endowed with its own base. Consider now a linear vector space $V$ with two bases, $e_1, e_2, \dots, e_n$ and $e'_1, e'_2, \dots, e'_n$. In what follows, they will be called the *old* base and the *new* base, respectively, reflecting our intention to change from the old base to the new base. Let $x = x_1e_1 + x_2e_2 + \dots + x_ne_n = x'_1e'_1 + x'_2e'_2 + \dots + x'_ne'_n$ be two coordinate decompositions of $x$ over the old and the new base. Assume that the vectors of the old base are decomposed by the new base

$$\begin{array}{ccccccccc} e_1 & = & c_{11}e'_1 & + & c_{21}e'_2 & + & \dots & + & c_{n1}e'_n, \\ e_2 & = & c_{12}e'_1 & + & c_{22}e'_2 & + & \dots & + & c_{n2}e'_n, \\ e_n & = & c_{1n}e'_1 & + & c_{2n}e'_2 & + & \dots & + & c_{nn}e'_n. \end{array}$$

Then, we get $x = x_1e_1 + x_2e_2 + \dots + x_ne_n = x_1(c_{11}e'_1 + c_{21}e'_2 + \dots + c_{n1}e'_n) + x_2(c_{12}e'_1 + c_{22}e'_2 + \dots + c_{n2}e'_n) + \dots + x_n(c_{1n}e'_1 + c_{2n}e'_2 + \dots + c_{nn}e'_n) = (x_1c_{11} + x_2c_{12} + \dots + x_nc_{1n})e'_1 + (x_1c_{21} + x_2c_{22} + \dots + x_nc_{2n})e'_2 + \dots + (x_1c_{n1} + x_2c_{n2} + \dots + x_nc_{nn})e'_n = x'_1e'_1 + x'_2e'_2 + \dots + x'_ne'_n$. Since the coordinates of a vector in a base are defined uniquely, we have

$$\left\{ \begin{array}{ccccccccc} x'_1 & = & c_{11}x_1 & + & c_{12}x_2 & + & \dots & + & c_{1n}x_n \\ x'_2 & = & c_{12}x_1 & + & c_{22}x_2 & + & \dots & + & c_{2n}x_n \\ x'_n & = & c_{1n}x_1 & + & c_{n2}x_2 & + & \dots & + & c_{nn}x_n \end{array} \right.$$

Thus, the change of bases incurs the change of coordinates, which has the form

$$x' = Cx,$$

[8]Can you rigorously show that the rotation angle if 120°?.

where $x'$ is the column-vector of coordinates in the new base, $x$ is the column-vector of coordinates in the old base, and the (square) matrix $C$ contains the coordinates of the old base vectors in the new base. The backward transformation of coordinates is expressed by using the inverse matrix, i.e.,

$$x = C^{-1}x'.$$

The matrix $C$ is called the *transition matrix* from the base $e_1, e_2, \ldots, e_n$ to the base $e'_1, e'_2, \ldots, e'_n$.

Consider now a linear operator $A : V \to V$ with the matrix $A$ in the old base (which plays the role of both $e_i$ and $f_i$ in the definition of a matrix of a linear operator) and with the matrix $A'$ in the new base. That is, $y = Ax$ and $y' = A'x'$. Besides that, $x' = Cx$ and $y' = Cy$ with the same transition matrix $C$. Thus, $y' = Cy = A'x' = A'Cx$, and then $y = C^{-1}A'Cx = Ax$. Since the matrix of a linear operator is defined uniquely by construction, we get

$$A = C^{-1}A'C, \text{ or } A' = CAC^{-1}.$$

This is the rule, by which the matrix of a linear operator is transformed under the change of base.

**Definition 16.** Two n×n matrices, $A$ and $B$, are called *conjugate* if there exists a non-degenerate matrix $C$ such that $B = CAC^{-1}$.

Thus, the matrix of a linear operator looks differently in different bases. Two matrices of the same linear operator in different bases are conjugate to each other. An obvious question is to what "standard" form can one transform the matrix of a linear operator by using conjugation. An incomplete, but still a reasonable answer to this question will be given in the next sections.

## 6.3 Eigenvectors and eigenvalues

**Definition 17.** Let $V$ be a vector space and let $A : V \to V$ be a linear operator. The vector $x \neq 0$ is said to be an *eigenvector* of $A$ with the *eigenvalue* $\lambda$ if $A(x) = \lambda x$. In other words, $A$ acts on $x$ as multiplication by scalars.

**Definition 18.** Let $A$ be a n×n matrix. The non-zero vector $x \in \mathbb{R}^n$ is called the *eigenvector* of $A$ with the *eigenvalue* $\lambda$ if $Ax = \lambda x$.

**Theorem 18.** *The number* $\lambda \in \mathbb{R}$ *is an eigenvalue of the matrix* $A$ *if and only if* $\det(A - \lambda E) = 0$.

*Proof.* Assume $\lambda$ is an eigenvalue of the matrix $A$. Then, for some $x \neq 0$ we have $Ax = \lambda x$, i.e.,

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & \lambda x_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & \lambda x_2 \\ & & & & \ldots & & & & \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \ldots & + & a_{nn}x_n & = & \lambda x_n \end{cases}$$

That is,

$$\begin{cases} (a_{11} - \lambda)x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & 0 \\ a_{21}x_1 & + & (a_{22} - \lambda)x_2 & + & \ldots & + & a_{2n}x_n & = & 0 \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \ldots & + & (a_{nn} - \lambda)x_n & = & 0 \end{cases},$$

which in matrix form is $(A-\lambda E)x = 0$. The latter is a homogeneous system of linear equations, which already has one solution $x = 0$. In order to have a non-zero solution, it must be degenerate (theorem 11), i.e., $\det(A - \lambda E) = 0$.

Conversely, if $\det(A - \lambda E) = 0$ then $(A - \lambda E)x = 0$ is degenerate and must have a non-zero solution $x$. Then, we have $Ax - \lambda x = 0$, concluding that $x$ is the eigenvector of $A$ with the eigenvalue $\lambda$. $\square$

Note that the final result of the decomposition of $\det(A - \lambda E)$ is a polynomial, i.e.,

$$\begin{vmatrix} (a_{11}-\lambda) & a_{12} & \dots & a_{1n} \\ a_{21} & (a_{22}-\lambda) & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & (a_{nn}-\lambda) \end{vmatrix} =$$

$= (a_{11} - \lambda)(a_{22} - \lambda)(a_{33} - \lambda)\dots(a_{nn} - \lambda) - a_{12}a_{21}(a_{33} - \lambda)\dots(a_{nn} - \lambda) + \dots = (-1)^n\lambda^n + c_{n-1}\lambda^{n-1} + c_{n-2}\lambda^{n-2} + c_1\lambda^1 + c_0$.

**Definition 19.** The polynomial function $f_A(\lambda) = \det(A - \lambda E)$ is called the *characteristic polynomial* of the matrix $A$.

**Theorem 19.** *Characteristic polynomials of conjugate matrices are equal.*

*Proof.* Let $B$ be the matrix that is conjugate to $A$, i.e., $B = CAC^{-1}$. Then, by the properties of matrix multiplication $f_B(\lambda) = \det(CAC^{-1} - \lambda E) = \det(CAC^{-1} - \lambda CEC^{-1}) = \det(C(A - \lambda E)C^{-1}) = \det(C)\det(A - \lambda E)\det(C^{-1}) = \det(A - \lambda E) = f_A(\lambda)$. $\square$

**Corollary 19.1.** The sum of diagonal elements of a matrix doesn't change under conjugation.

*Proof.* Note that the entire characteristic polynomial doesn't change under conjugation. In particular, the factor $c_{n-1}$ at $\lambda^{n-1}$ doesn't change, too. In fact, the product of $n-1$ copies of $\lambda$ can only appear from the term $(a_{11} - \lambda)(a_{22} - \lambda)(a_{33} - \lambda)\dots(a_{nn} - \lambda)$ because the other terms contain $\lambda$ to the power of $n - 2$ or less. Then, $c_{n-1}\lambda^{n-1} = (-1)^{n-1}\lambda^{n-1}(a_{11} + a_{22} + \dots + a_{nn})$. Thus, the sum of diagonal elements of a matrix doesn't change under conjugation. $\square$

**Definition 20.** The sum of matrix's diagonal elements is called its *trace* and is denoted by $\mathrm{tr}(A)$.

**Example 6.4.** Find eigenvectors and eigenvalues of the matrix

$$A = \begin{pmatrix} 3 & 4 \\ 5 & 2 \end{pmatrix}.$$

**Solution.** $A - \lambda E = \begin{pmatrix} 3 & 4 \\ 5 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 5 & 2 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} 3-\lambda & 4 \\ 5 & 2-\lambda \end{pmatrix}$.

Thus, we get $\begin{vmatrix} 3-\lambda & 4 \\ 5 & 2-\lambda \end{vmatrix} = (3-\lambda)(2-\lambda) - 20 = \lambda^2 - 5\lambda - 14 = 0$. Solving the quadratic equation, we get $\lambda_1 = -2$ and $\lambda_2 = 7$.

Case 1: $\lambda_1 = 7$.

$$A - 7E = \begin{pmatrix} -4 & 4 \\ 5 & -5 \end{pmatrix}$$

$$\begin{pmatrix} -4 & 4 \\ 5 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\left( \begin{array}{cc|c} -4 & 4 & 0 \\ 5 & -5 & 0 \end{array} \right) \xrightarrow{[1]\mapsto[1]/(-4)} \left( \begin{array}{cc|c} 1 & -1 & 0 \\ 5 & -5 & 0 \end{array} \right) \xrightarrow{[2]\mapsto[2]-[1]\times 5} \left( \begin{array}{cc|c} 1 & -1 & 0 \end{array} \right)$$

There exists a non-zero solution, i.e., the vector $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (and all the vectors that are proportionate to it) are eigenvectors with the eigenvalue $\lambda_1 = 7$.

Case 2: $\lambda_2 = 2$.

$$A - \lambda E = \begin{pmatrix} 5 & 4 \\ 5 & 4 \end{pmatrix}$$

Similarly, we get

$$\begin{pmatrix} 5 & 4 \\ 5 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and the vector $y = \begin{pmatrix} 4 \\ -5 \end{pmatrix}$ is the eigenvector with the eigenvalue $\lambda_2 = 2$.

**Example 6.5.** Find eigenvectors and eigenvalues of the matrix

$$\begin{pmatrix} 5 & -7 & 7 \\ -8 & 3 & -2 \\ -8 & -1 & 2 \end{pmatrix}$$

**Solution.**

$$\det \begin{pmatrix} 5-\lambda & -7 & 7 \\ -8 & 3-\lambda & -2 \\ -8 & -1 & 2-\lambda \end{pmatrix} =$$

$= (5-\lambda)\times(3-\lambda)\times(2-\lambda)-7\times 8\times 2+8\times 1\times 7+8\times 7\times(3-\lambda)-8\times 7\times(2-\lambda)-1\times 2\times(5-\lambda) = (5-\lambda)\times(1-\lambda)\times(4-\lambda) = 0$.

Case 1: $\lambda_1 = 5$.

$$A - \lambda E = \begin{pmatrix} 0 & -7 & 7 \\ -8 & -2 & -2 \\ -8 & -1 & -3 \end{pmatrix}$$

$$\begin{pmatrix} 0 & -7 & 7 \\ -8 & -2 & -2 \\ -8 & -1 & -3 \end{pmatrix} \to \begin{pmatrix} 0 & 1 & -1 \\ -8 & -2 & -2 \\ -8 & -1 & -3 \end{pmatrix} \to \begin{pmatrix} 0 & 1 & -1 \\ -8 & 0 & -4 \\ -8 & 0 & -4 \end{pmatrix} \to \begin{pmatrix} 0 & 1 & -1 \\ 2 & 0 & 1 \end{pmatrix}$$

Then, letting $x_3 = -2$ we get $x_2 = -2$ and $x_1 = 1$. The first eigenvector is $x = \begin{pmatrix} 1 \\ -2 \\ -2 \end{pmatrix}$.

Case 2: $\lambda_1 = 1$.

$$A - \lambda E = \begin{pmatrix} 4 & -7 & 7 \\ -8 & 2 & -2 \\ -8 & -1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 4 & -7 & 7 \\ -8 & 2 & -2 \\ -8 & -1 & 1 \end{pmatrix} \to \begin{pmatrix} 4 & -7 & 7 \\ -8 & 2 & -2 \\ 1 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 0 & -7 & 7 \\ 0 & 1 & -1 \\ 1 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}.$$

Letting $x_3 = 1$ we get $x_2 = 1$ and $x_1 = 0$. The second eigenvector is $y = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.

Case 3: $\lambda_1 = 4$.

$$A - \lambda E = \begin{pmatrix} 1 & -7 & 7 \\ -8 & -1 & -2 \\ -8 & -1 & -2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -7 & 7 \\ -8 & -1 & -2 \\ -8 & -1 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & -7 & 7 \\ 8 & 1 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & -7 & 7 \\ 0 & -57 & 54 \end{pmatrix} \to \begin{pmatrix} 1 & -7 & 7 \\ 0 & 19 & -18 \end{pmatrix}$$

Guessing that $x_3 = 19$, we get $x_2 = 18$ and $x_1 = -7$. Therefore, the last eigenvector is $z = \begin{pmatrix} -7 \\ 18 \\ 19 \end{pmatrix}$. □

In the example above, one could probably notice that the transition matrix

$$C = \begin{pmatrix} 1 & 0 & -7 \\ -2 & 1 & 18 \\ -2 & 1 & 19 \end{pmatrix},$$

which is obtained from the vectors $x$, $y$, and $z$ by attaching their coordinates to each other, transforms matrix $A$ by the rule $C^{-1}AC$ to the diagonal matrix

$$\begin{pmatrix} 1 & -7 & 7 \\ 2 & 5 & -4 \\ 0 & -1 & 1 \end{pmatrix} \times \begin{pmatrix} 5 & -7 & 7 \\ -8 & 3 & -2 \\ -8 & -1 & 2 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & -7 \\ -2 & 1 & 18 \\ -2 & 1 & 19 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix},$$

where the diagonal entries are the eigenvalues of $A$. It happens because $x$, $y$, and $z$ are linearly independent and form a base in $\mathbb{R}^3$, while the matrix of a linear operator in a base that consists of eigenvectors must always be diagonal (by definition).

**Theorem 20.** *Let $A$ be a $n \times n$ matrix. Eigenvectors of $A$ corresponding to different eigenvalues are linearly independent.*

*Proof.* Let $x_1, x_2, \ldots, x_k$ be the eigenvectors corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$. We need to prove that if $\forall i \neq j\ \lambda_i \neq \lambda_j$ then $x_1, x_2, \ldots, x_k$ are linearly independent.

The proof of this statement is by induction on $k$. The base of the induction ($k = 1$) is obvious because every eigenvector is non-zero by definition and, therefore, forms a linearly independent set by itself. Assume that we have proven the statement for every set of $k-1$ eigenvectors with different eigenvalues, and now want to proceed to $k$ eigenvectors.

Suppose, on the contrary, that $x_1, x_2, \ldots, x_k$ are linearly dependent, i.e., there exist coefficients $\mu_1, \mu_2, \ldots, \mu_k$, some of which are non-zero, such that

$$\mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_n x_n = 0.$$

By applying $A$ to both sides of the equation, we get $A(\mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_n x_n) = A(\mu_1 x_1) + A(\mu_2 x_2) + \cdots + A(\mu_n x_n) = \mu_1 A(x_1) + \mu_2 A(x_2) + \cdots + \mu_n A(x_n) = \mu_1 \lambda_1 x_1 + \mu_2 \lambda_2 x_2 + \cdots + \mu_n \lambda_n x_n$. This, in combination with the linear combination of $x_1, x_2, \ldots, x_k$ multiplied by $\lambda_1$, gives

$$\begin{cases} \mu_1 \lambda_1 x_1 + \mu_2 \lambda_2 x_2 + \cdots + \mu_n \lambda_n x_n = 0 \\ \mu_1 \lambda_1 x_1 + \mu_2 \lambda_1 x_2 + \cdots + \mu_n \lambda_1 x_n = 0 \end{cases},$$

which implies that

$$\mu_2(\lambda_2 - \lambda_1)x_2 + \mu_3(\lambda_3 - \lambda_1)x_3 + \cdots + \mu_n(\lambda_n - \lambda_1)x_n = 0.$$

Thus, we have a linear combination of $x_2, \ldots, x_n$ which consists of $k-1$ vectors and is linearly dependent. By induction, it is possible only if such linear combination is trivial, i.e., if $\mu_2(\lambda_2 - \lambda_1) = \mu_3(\lambda_3 - \lambda_1) = \cdots = \mu_n(\lambda_n - \lambda_1) = 0$. Since all $\lambda$'s are different, we have $\lambda_i - \lambda_1 \neq 0$ and, therefore, $\mu_2 = \mu_3 = \cdots = \mu_n = 0$.

Repeating the same argument with multiplication of the linear combination of $x_1, x_2, \ldots, x_k$ by $\lambda_2$ instead of $\lambda_1$ leads to the conclusion that $\mu_1 = 0$, too. Therefore, $\mu_1 = \mu_2 = \cdots = \mu_n = 0$ and the linear combination we chose is trivial, i.e., $x_1, x_2, \ldots, x_k$ are linearly independent. □

## 6.4 Diagonalizable matrices

The property of a linear operator to have a base consisting of eigenvectors (example 6.5) is very useful. Such "good" linear operators play an important role in many applications, for instance, in linear ordinary differential equations and in linear differential-difference equations.

**Definition 21.** The linear operator $A : V \to V$ is called *diagonalizable* if its matrix is diagonal in some base of $V$. Equivalently, one can find a base of $V$ which consists of eigenvectors of $A$.

**Definition 22.** A n×n matrix $A$ is called *diagonalizable*[9] if there exist a non-degenerate matrix $C$ and a diagonal matrix $D$ such that $CAC^{-1} = D$.

Obviously, a linear operator is diagonalizable if and only if its matrix (in some base) is diagonalizable. Considering that the characteristic polynomial of $f_A(\lambda)$ has the degree $n$ and, therefore, can have at most $n$ roots, the following statement follows immediately from theorem 20.

[9] Note that the statements $CAC^{-1} = D$ and $C^{-1}AC = D$ are equivalent up to the difference between $C$ and $C^{-1}$. Since the transition matrix $C$ is not specified in this definition, we ignore this difference from now on.

**Theorem 21.** *A n×n matrix which has n distinct real eigenvalues is diagonalizable.*

However, this sufficient condition is not necessary, i.e., there exist diagonalizable matrices with coinciding eigenvalues.

**Example 6.6.**

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

$\det(A - \lambda E) = (1 - \lambda)^3$, i.e., $\lambda = 1$. However, $A - \lambda E = 0$, so every vector in $\mathbb{R}^3$ is an eigenvector for $A$, i.e., one can find a base consisting of eigenvectors of $A$. That is, $A$ is diagonalizable (already in the diagonal form).

**Example 6.7.** The matrix $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ is not diagonalizable. Indeed, $\det(A - \lambda E) = (2 - \lambda)^3$, i.e., $\lambda = 2$. Then, all eigenvectors of $A$ must satisfy the SLE given by

$$\left(\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right).$$

However, all its solutions are proportional to the vector $x = (1, 0, 0)$ and thus cannot form a base in $\mathbb{R}^3$. Therefore, $A$ is not diagonalizable.

The problem with the matrix $A$ in the example 6.7 is that the multiplicity of the root $\lambda$ is equal to three, i.e., the characteristic polynomial contains three copies of the corresponding linear term, while the corresponding SLE has only one-dimensional subspace of eigenvectors. In example 6.6, the multiplicity of the root was also equal to three, but the corresponding SLE had a three-dimensional space of solutions.

For each eigenvalue $\lambda_0$ of a matrix, there are two numbers measuring, roughly speaking, the number of eigenvectors corresponding to $\lambda_0$. These numbers are called multiplicities of $\lambda_0$.

**Definition 23.** Let $\lambda_0$ be a root of the characteristic polynomial for an n×n matrix $A$. The *algebraic multiplicity* of $\lambda_0$ is the highest power $k$ such that $(\lambda - \lambda_0)^k$ is a factor of $f_A(\lambda)$. The *geometric multiplicity* of $\lambda_0$ is the number of linearly independent solutions to $(A - \lambda_0 E)x = 0$, i.e., the geometric multiplicity is equal to $n - \mathrm{rk}(A - \lambda_0 E)$.

**Lemma 22.** *The algebraic multiplicity of a root is greater or equal to its geometric multiplicity.*

**Theorem 23.** *A n×n matrix is diagonalizable if and only if all roots of $f_A(\lambda)$ are real and their algebraic multiplicities are equal to their geometric multiplicities.*

*Proof* of these two statements falls beyond the scope of this chapter. □

Diagonalizable matrixes are very special because they are conjugate to diagonal matrices. Geometrically, a diagonalizable matrix is an inhomogeneous dilation: it scales the vector space, as does a homogeneous dilation, but by a different factor in each dimension, determined by the scale factors on each axis (diagonal entries).

Diagonalization can be used to compute the powers of a matrix $A$, provided that the matrix is diagonalizable. Suppose we have found that

$$A = CDC^{-1},$$

and we need to compute $A^k$. Diagonal matrices are especially easy to handle: their eigenvalues and eigenvectors are known and one can raise a diagonal matrix to a power by simply raising the diagonal entries to the same power, i.e., if $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ then $D^k = \mathrm{diag}(\lambda_1^k, \ldots, \lambda_n^k)$. Then,

$$A^k = CD(C^{-1}C)D(C^{-1}\ldots C)DC^{-1} = CD^kC^{-1}.$$

That is, in order to power a diagonalizable matrix, one needs to transform it to the diagonal form, take the desired power of the diagonal form, and transform back by using the same transition matrix.

**Example 6.8.** Compute $A^{20}$ for $A = \begin{pmatrix} 5 & -5 & -8 \\ -8 & 2 & 8 \\ -1 & -5 & -2 \end{pmatrix}$.

**Solution.** First, since $A$ has three distinct eigenvalues ($\lambda_1 = -3$, $\lambda_2 = 2$, $\lambda_3 = 6$), it is diagonalizable. Next, we find the diagonalization

$$\begin{pmatrix} 5 & -5 & -8 \\ -8 & 2 & 8 \\ -1 & -5 & -2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1 & 2 \\ 1 & -1 & -1 \end{pmatrix} \times \begin{pmatrix} -3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix} \times \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{pmatrix}.$$

Finally,

$$A^{20} = \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1 & 2 \\ 1 & -1 & -1 \end{pmatrix} \times \begin{pmatrix} 3^{20} & 0 & 0 \\ 0 & 2^{20} & 0 \\ 0 & 0 & 6^{20} \end{pmatrix} \times \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{pmatrix}.$$

Computation of matrix's power can be extended to define a *function* of a matrix for every analytical real function[10]. Let

$$\varphi(x) = \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(0)}{k!} x^k$$

be such a function. Then, having defined the convergence of matrix series, we can define

$$\varphi(A) = \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(0)}{k!} A^k,$$

[10] A function is called analytical if it is equal to its (absolutely converging) Mc'Laurin series.

since matrix addition and matrix multiplication are correctly defined. Additionally, if $A$ is diagonalizable then $A = CDC^{-1}$ and

$$\varphi(A) = \varphi(CDC^{-1}) = \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(0)}{k!}(CDC^{-1})^k = C\left(\sum_{k=0}^{\infty} \frac{\varphi^{(k)}(0)}{k!} D^k\right) C^{-1}.$$

If $D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ then $\varphi(D) = \operatorname{diag}(\varphi(\lambda_1), \varphi(\lambda_2), \ldots, \varphi(\lambda_n))$ and

$$\varphi(A) = C \times \operatorname{diag}(\varphi(\lambda_1), \varphi(\lambda_2), \ldots, \varphi(\lambda_n)) \times C^{-1}.$$

**Example 6.9.** Compute $\exp(A)$ for $A = \begin{pmatrix} -7 & -8 & -8 \\ 4 & 5 & 4 \\ 2 & 4 & -1 \end{pmatrix}$.

**Solution.** First, we compute eigenvalues $\lambda_1 = -1$, $\lambda_2 = -3$, $\lambda_3 = 1$ and note that $A$ is diagonalizable because it has three distinct eigenvalues. Next,

$$A = \begin{pmatrix} -4 & -2 & 3 \\ 2 & 1 & -2 \\ 1 & 0 & -1 \end{pmatrix} \times \begin{pmatrix} -1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} -1 & -2 & 1 \\ 0 & 1 & -2 \\ -1 & -2 & 0 \end{pmatrix}, \quad \text{and}$$

$$\exp(A) = \begin{pmatrix} -4 & -2 & 3 \\ 2 & 1 & -2 \\ 1 & 0 & -1 \end{pmatrix} \times \begin{pmatrix} \frac{1}{e} & 0 & 0 \\ 0 & \frac{1}{e^3} & 0 \\ 0 & 0 & e \end{pmatrix} \times \begin{pmatrix} -1 & -2 & 1 \\ 0 & 1 & -2 \\ -1 & -2 & 0 \end{pmatrix}.$$

## 6.5 Exercises

Problem 6.1. Find eigenvalues of the following matrices

(a) $\begin{pmatrix} 5 & -2 \\ 1 & 8 \end{pmatrix}$ (b) $\begin{pmatrix} 7 & 6 \\ -8 & -7 \end{pmatrix}$ (c) $\begin{pmatrix} 4 & -8 \\ 4 & -8 \end{pmatrix}$ (d) $\begin{pmatrix} -8 & -1 \\ 6 & -3 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & -7 & -7 \\ 7 & -5 & 1 \\ -7 & 7 & 1 \end{pmatrix}$ (f) $\begin{pmatrix} -6 & 2 & -4 \\ -8 & -3 & 2 \\ -4 & 2 & -6 \end{pmatrix}$ (g) $\begin{pmatrix} -5 & -2 & -2 \\ -8 & 1 & -2 \\ 8 & -8 & -5 \end{pmatrix}$

(h) $\begin{pmatrix} -4 & -4 & -4 & -4 \\ 0 & -2 & 5 & 5 \\ 2 & 4 & -7 & -5 \\ -2 & -4 & 6 & 4 \end{pmatrix}$ (i) $\begin{pmatrix} -8 & 7 & -1 & -3 \\ -2 & 1 & 2 & -4 \\ 4 & -4 & 1 & -1 \\ 4 & -4 & 6 & -6 \end{pmatrix}$ (j) $\begin{pmatrix} 6 & 0 & -4 \\ 24 & 6 & -44 \\ 8 & 0 & -6 \end{pmatrix}$

Problem 6.2. Find eigenvalues and eigenvectors of the following matrices and perform diagonalization, when possible.

(a) $\begin{pmatrix} 3 & -6 \\ 4 & -7 \end{pmatrix}$ (b) $\begin{pmatrix} 5 & 4 \\ 0 & 7 \end{pmatrix}$ (c) $\begin{pmatrix} -6 & -2 \\ 3 & -1 \end{pmatrix}$ (d) $\begin{pmatrix} 6 & 14 \\ 0 & -8 \end{pmatrix}$

(e) $\begin{pmatrix} -2 & -6 & -2 \\ 2 & 2 & -2 \\ -2 & -6 & -2 \end{pmatrix}$ (f) $\begin{pmatrix} 3 & -8 & 3 \\ -3 & -2 & 3 \\ 2 & -8 & 4 \end{pmatrix}$ (g) $\begin{pmatrix} -7 & 0 & -10 \\ 0 & 7 & 4 \\ 0 & 0 & 3 \end{pmatrix}$

(h) $\begin{pmatrix} -2 & 5 & -8 \\ 2 & 1 & -8 \\ -6 & 6 & -1 \end{pmatrix}$ (i) $\begin{pmatrix} -4 & 2 & -2 \\ -1 & -7 & 2 \\ -1 & -2 & -3 \end{pmatrix}$ (j) $\begin{pmatrix} 3 & 7 & -3 \\ 3 & -1 & 3 \\ -5 & -7 & 1 \end{pmatrix}$

(k) $\begin{pmatrix} 3 & -2 & 3 & -2 \\ -2 & 6 & 0 & 2 \\ 2 & -8 & -2 & -2 \\ 3 & -6 & 3 & -2 \end{pmatrix}$ (l) $\begin{pmatrix} 0 & 6 & 1 & 8 \\ -8 & -8 & 1 & 0 \\ -6 & 0 & -7 & -2 \\ 3 & 3 & 4 & 5 \end{pmatrix}$

# 7 Quadratic forms

Let $x_1, x_2, \dots, x_n$ be a set of variables. A *monomial* is a product of non-negative integer powers of $x_1, x_2, \dots, x_n$. For instance, $x_1^2x_2x_3^4$ is a monomial. The *degree* of a monomial with respect to the variable $x_i$ is the power, at which $x_i$ occurs in it. For instance, the degree of $x_1^2x_2x_3^4$ with respect to $x_1$ is equal to two. The *total degree* of the monomial is the sum of its degrees with respect to all the variables. For instance, the total degree of $x_1^2x_2x_3^4$ is equal to seven.

A *homogeneous polynomial* is the sum of monomials, whose terms have the same total degree. For example, $x_1^5 + 2x_1^3x_2^2 + 9x_2^5$ is a homogeneous polynomial of degree five. The polynomial $x_1^3 + 3x_1^2x_2^3 + x_2^7$ is not homogeneous, because the sum of exponents does not match from term to term.

A *quadratic form* is a homogeneous polynomial of degree two in $n$ variables. Quadratic forms play an important role in many branches of mathematics, including calculus and optimization theory.

## 7.1 Matrix of a quadratic form

Each quadratic form in $n$ variables can be written in a canonical form which involves matrix notation. For instance, $Q(x_1, x_2) = x_1^2 + 6x_1x_2 + 5x_2^2$ can be written as matrix product

$$(x_1, x_2)\begin{pmatrix} 1 & 3 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1, x_2)\begin{pmatrix} x_1 + 3x_2 \\ 3x_1 + 5x_2 \end{pmatrix} = x_1^2 + 6x_1x_2 + 5x_2^2.$$

Note that the matrix product contains two copies of the monomial $3x_1x_2$, which sum up to $6x_1x_2$. In general, the matrix product of this form can be transformed as

$$(x_1x_2\dots x_n)\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} =$$

$$= (x_1x_2\dots x_n)\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{pmatrix} =$$

$= x_1(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}) + x_2(a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n) + \cdots + x_n(a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n) = a_{11}x_1^2 + (a_{12} + a_{21})x_1x_2 + \cdots + a_{22}x_2^2 + \cdots + (a_{1n} + a_{n1})x_1x_n + \cdots + a_{nn}x_n^2$. Since the coefficients $a_{ij}$ are yet to be defined, it is convenient to assume that $a_{ij} = a_{ji}$ and to include the factor of 2 in the general expression of a quadratic form:

$$Q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j.$$

**Definition 24.** A n×n matrix $A = (a_{ij})$ is called *symmetric* (*anti-symmetric*) if $\forall i, j\ a_{ij} = a_{ji}$ ($\forall i, j\ a_{ij} = -a_{ji}$). In other words, $A$ is symmetric (anti-symmetric) if $A = A^T$ ($A = -A^T$).

Thus, every symmetric matrix defines a quadratic form and vice versa, every quadratic form can be defined by a symmetric matrix by the identity

$$Q(x) = x^T Ax,$$

where $x = (x_1x_2 \ldots x_n)^T$ is a column vector. As a rule, one must put the full square coefficients in the diagonal of the matrix, while the coefficients at the cross-product terms are divided by two and placed in the corresponding cells, one above and one below the diagonal.

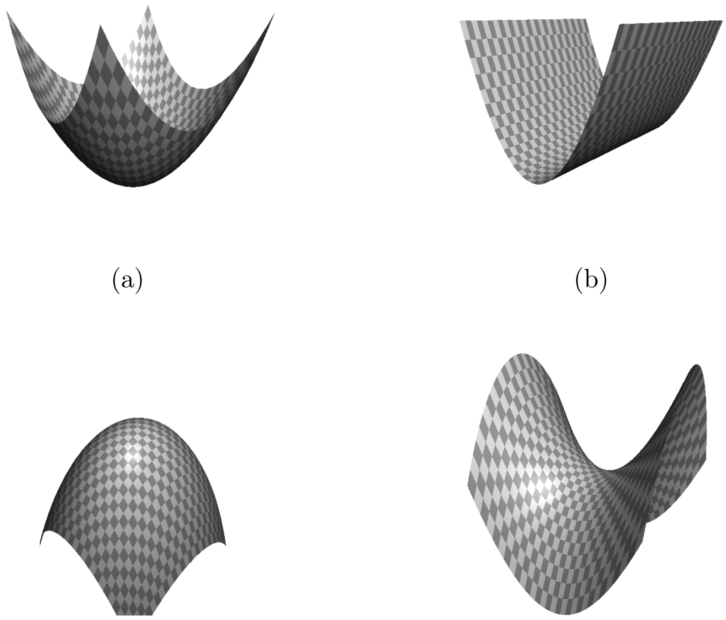
**Example 7.1.** $Q(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + 3x_3^2 + 4x_1x_3 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 3 \end{pmatrix}$

**Example 7.2.** $Q(x_1, x_2, x_3) = 4x_1x_2 + 2x_1x_3 + 10x_2x_3 = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 5 \\ 1 & 5 & 0 \end{pmatrix}$

**Definition 25.** A quadratic form $Q(x_1, x_2, \ldots, x_n)$ is called *positive definite*[11] if $Q(x_1, x_2, \ldots, x_n) > 0$ for every non-zero vector $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Similarly, $Q(x_1, x_2, \ldots, x_n)$ is called *negative definite* (*positive semidefinite*, *negative semidefinite*) if $Q(x_1, x_2, \ldots, x_n) < 0$ ($Q(x_1, x_2, \ldots, x_n) \geq 0$, $Q(x_1, x_2, \ldots, x_n) \leq 0$, respectively). Otherwise, a quadratic form is called *not definite.*

Figure 2 gives a simple and intuitive explanation of what is a positive definite quadratic form. First of all, if $z = f(x, y)$ is a quadratic form then its planar sections $z(y) = f(0, y)$ and $z(x) = f(x, 0)$ both have to be quadratic functions.

For instance, the 3D-graph of $z = x^2 + y^2$ shown in part (a) is a surface with concave up planar cross-sections for both $x = 0$ ($z = y^2$) and $y = 0$ ($z = x^2$). The quadratic form $z = x^2+y^2$ is positive definite because it takes only positive values for non-zero pairs $(x, y)$ and is equal to zero if and only if both $x$ and $y$ are equal to zero. Now, if we don't require $z$ be non-zero for every non-zero pair $(x, y)$ then we get a non-positive definite quadratic form whose surface is shown in part (b). An example of such function would be the function $z = x^2 - 2xy + y^2 = (x - y)^2 \geq 0$, which is zero for $x = y = 1$. Similarly, the negative-definite quadratic function $z = -x^2 - y^2$ has planar cross-sections for both $x = 0$ ($z = -y^2$) and $y = 0$ ($z = -x^2$) but this time the cross-sections are concave down (figure 2 (c)). A non-definite quadratic form is such that some of its



(a) (b)

(c) (d)

Figure 2: 3D shapes of (a) positive definite, (b) non-negative definite, (c) negative-definite, and (d) not definite quadratic functions.

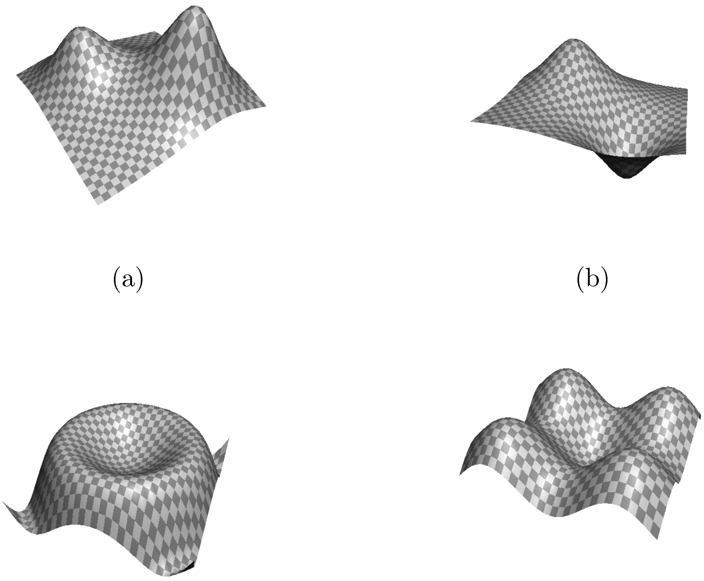
cross-sections are are concave up, while others are concave down. For instance, the quadratic form $z = x^2 + y^2$ is non-definite as $z = x^2$ for $y = 0$ and $z = -y^2$ for $x = 0$.

One characteristic spot on the surface of a non-definite quadratic function is called *saddle point*. Saddle point has zero slope, but the surface is concave up in one direction and concave down in the other direction. It is the point located right in the middle of the surface shown in figure 2 (d). One can think of a saddle point as of unstable equilibrium, where it is a local minimum in one cross-section but at the same time it is a local maximum in another cross-section. Generic surfaces may have various numbers of maxima, minima, and saddle points, and quadratic forms are often used as a tool to approximate the local behavior of generic surfaces. For instance, the surface shown in figure 7.1(a) can be described as two hills (or mountains) and two respective valleys. The 'best' path from one valley to the other passes through the saddle point. Saddle points play an important role in multi-dimensional calculus and in the theory of ordinary differential equations.

## 7.2 Change of base

We have seen in section 6.2 that the matrix of a linear operator changes according to the rule $A \mapsto CAC^{-1}$ under the linear transformation of the vector space induced by the matrix $C$. In this section we investigate how does the matrix of a quadratic form change under a similar

[11] Do not confuse the adjective *définite* with the past participle *defíned*.



(a) (b)

(c) (d)

Figure 3: Examples of extreme points on generic surfaces. (a) two local maxima separated by a saddle-point, (b) a local maximum and a local minimum, no saddle point in between, (c) one local minimum and degenerate local maxima and minima, (d) four local maxima, one local minimum, and four saddle points.

transformation.

Consider two bases, $e_1, e_2, \ldots, e_n$ and $e'_1, e'_2, \ldots, e'_n$ of the vector space $V$ and let $x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n = x'_1 e'_1 + x'_2 e'_2 + \cdots + x'_n e'_n$ be two coordinate decompositions of the vector $x$ over these two bases. In section 6.2 we saw that the new coordinates are expressed through the old coordinates as

$$\left\{ \begin{array}{ccccccccc} x'_1 & = & c_{11}x_1 & + & c_{12}x_2 & + & \ldots & + & c_{1n}x_n \\ x'_2 & = & c_{12}x_1 & + & c_{22}x_2 & + & \ldots & + & c_{2n}x_n \\ x'_n & = & c_{1n}x_1 & + & c_{n2}x_2 & + & \ldots & + & c_{nn}x_n \end{array} \right. ,$$

where matrix $C$ consists of the coordinates of old base vectors in the new base, i.e., $x' = Cx$.

Let $A$ and $A'$ be the matrices of a quadratic form in the old and new base, respectively. We have

$$Q(x) = (x')^T A' x' = (Cx)^T A' C x = x^T (C^T A' C) x = x^T A x.$$

Therefore, the matrix of a quadratic form transforms as $A \mapsto C^T A C$, where $C$ is a transition matrix[12]. In the next section we will ask to what standard form can one transform a quadratic form by a change of base.

[12]Perhaps, we should be more careful with notation because the matrix $C$ here is inverse to the matrix $C$ discussed in the previous paragraph. However, since $C$ is unknown anyway, we will keep using $C$ instead of $C^{-1}$ for simplicity of notation.

## 7.3 Canonical diagonal form

**Theorem 24.** *Let $Q(x)$ be a quadratic form in $n$ variables. There exists a change of base such that in the new base*

$$Q(x) = x_1^2 + x_2^2 + \cdots + x_k^2 - x_{k+1}^2 - x_{k+2}^2 - \cdots - x_{k+l}^2.$$

*Proof.* Assume that $Q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j$ and suppose that $a_{11} \neq 0$. We will look for the change of coordinates of the form $x_1 = p_1y_1 + p_2y_2 + \cdots + p_ny_n$ and $x_i = y_i$ for $i \geq 2$. We get $Q(x_1, x_2, \ldots, x_n) = a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n + \sum_{i=2}^{n} a_{ii}x_i^2 + 2\sum_{2\leq i<j} a_{ij}x_ix_j = a_{11}x_1^2 + 2x_1(a_{12}x_2 + \cdots + a_{1n}x_n) + R(x_2, \ldots, x_n)$, where $R(x_2, \ldots, x_n)$ is a quadratic form in the $n-1$ variables.

Then, $Q(y) = a_{11}(p_1y_1 + p_2y_2 + \cdots + p_ny_n)^2 + 2p_1y_1(a_{12}y_2 + \cdots + a_{1n}y_n) + R_1(y_2, \ldots, y_n)$, where the terms containing $y_2, \ldots, y_n$ went to $R_1(y_2, \ldots, y_n)$. Then, $Q(y) = a_{11}(p_1^2y_1^2 + 2p_1p_2y_1y_2 + 2p_1p_3y_1y_3 + \cdots + 2p_1p_ny_1y_n) + 2(p_1a_{12}y_1y_2 + p_1a_{13}y_1y_3 + \cdots + p_1a_{1n}y_1y_n) + R_2(y_2, \ldots, y_n)$, where $R_2(y_2, \ldots, y_n)$ contains all the terms with $y_2, \ldots, y_n$. If we set $p_2 = -a_{12}, p_3 = -a_{13}, \ldots, p_n = -a_{1n}$ and $p_1 = \sqrt{|a_{11}|}$ then $Q(y) = \pm y_1^2 + R_1(y_2, \ldots, y_n)$, where the sign at $y_1^2$ is the same as the sign of $a_{11}$.

Now consider the case when $a_{11} = 0$ but at least one of $a_{1i} \neq 0$. Without loss of generality, assume that $a_{12} \neq 0$. Consider the change of coordinates of the form $x_1 = y_1 + y_2$, $x_2 = y_1 - y_2$, and $x_i = y_i$ for $i \geq 3$. We are now back to the case when $a_{11} \neq 0$ because $Q(y) = (y_1 + y_2)(y_1 - y_2) + \cdots = y_1^2 - y_2^2 + \ldots$ contains $y_1^2$.

Finally, if $a_{1i} = 0$ for all $i$, we can proceed to a quadratic form with fewer number of variables. □

The pair of numbers $(k, l)$, which are the number of positive and negative terms in the canonical representation, respectively, is defined uniquely and will be called *the signature* of the quadratic form. When there are no zero terms in the canonical form (i.e., $k + l = n$) and the dimension $n$ of the enveloping vector space is known, it is enough to know the difference $k - l$, which is sometimes also called signature[13]. For instance, the signature $(3, 1)$ in $\mathbb{R}^4$ (also known as Minkowski space) is also called "signature 2". We will no longer pay any attention to the distinction between these two formulations.

The proof of this theorem contains the procedure called the *completion of full squares*, which can be used to find the canonical base.

**Example 7.3.** Find the canonical form of

$$Q(x_1, x_2, x_3) = x_1^2 - 4x_1x_2 + 2x_1x_3 + 6x_2^2 + 2x_3^2$$

[13] This second notation is found frequently in physics books

**Solution.** $Q(x_1, x_2, x_3) = (x_1^2 - 4x_1x_2 + 4x_2^2) + 2x_1x_3 + 2x_2^2 + 2x_3^2 = (x_1 - 2x_2)^2 + 2x_1x_3 + 2x_2^2 + 2x_3^2$. Denote $y_1 = x_1 - 2x_2$, $y_2 = x_2$, and $y_3 = x_3$. Then, $x_1 = y_1 + 2y_2$ and we get $Q(y_1, y_2, y_3) = y_1^2 + 2(y_1 + 2y_2)y_3 + 2y_2^2 + 2y_3^2 = y_1^2 + 2y_1y_3 + 4y_2y_3 + 2y_2^2 + 2y_3^2 = (y_1^2 + 2y_1y_3 + y_3^2) + 4y_2y_3 + 2y_2^2 + y_3^2 = (y_1 + y_3)^2 + 4y_2y_3 + 2y_2^2 + y_3^2$. Now denote $y_1 + y_3 = z_1$, $y_2 = z_2$, and $y_3 = z_3$. We get $Q(z_1, z_2, z_3) = z_1^2 + 2z_2^2 + 4z_2z_3 + z_3^2 = z_1^2 + 2(z_2^2 + 2z_2z_3 + z_3^2) - z_3^2 = z_1^2 + 2(z_2 + z_3)^2 - z_3^2$. Finally, denote $t_1 = z_1$, $t_2 = z_2 + z_3$, and $t_3 = z_3$. We get $Q(z_1, z_2, z_3) = t_1^2 + 2t_2^2 - t_3^2$. Thus, the signature of this quadratic form is $(2, 1)$.

**Theorem 25.** *A quadratic form $Q(x_1, x_2, \ldots, x_n)$ is positive definite (negative definite, positive semidefinite, negative semidefinite) if and only if its signature is $(n, 0)$ (its signature is $(0, n)$, $(k, 0)$, and $(0, k)$, respectively, where $k < n$). A quadratic form $Q(x_1, x_2, \ldots, x_n)$ is not definite if and only if its signature is $(k, l)$, where $k > 0$ and $l > 0$.*

*Proof follows immediately from the definition.* □

## 7.4 Sylvester's criterion

Sometimes the completion of least squares is not the most convenient way to check whether or not the quadratic form is positive definite. The following theorem helps to compute the signature of a quadratic form without cumbersome coordinate changes.

**Theorem 26** (Sylvester's criterion)**.** *Let $Q(x) = x^TAx$ be a quadratic form in $n$ variables where $A = A^T$. The form $Q(x)$ is positive definite if and only if all upper left corner $k \times k$ minors of $A$ are positive, i.e.,*

$$\Delta_1 = a_{11} > 0,$$

$$\Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0,$$

$$\Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \ldots$$

$$\Delta_n = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} > 0.$$

This theorem follows from the following lemma.

**Lemma 27.** *Let $x_1 = p_1y_1 + p_2y_2 + \cdots + p_ny_n$ and $x_i = y_i$ for $i \geq 2$ be the change of variables in the proof of theorem 24. All upper-left corner $k \times k$ minors of $A$ do not change if $p_1 = 1$.*

**Corollary 27.1.** Let $Q(x) = x^TAx$ be a quadratic form in $n$ variables and let $\lambda_1x_1^2 + \lambda_2x_2^2 + \cdots + \lambda_nx_n^2$ be its diagonal representation obtained by a series of full square completion steps of the form $x_1 = p_1y_1 + p_2y_2 + \cdots + p_ny_n$, $x_i = y_i$ for $i \geq 2$ with $p_1 = 1$. Assume that $\Delta_i \neq 0$ for all $i$. Then,

$$\lambda_i = \frac{\Delta_i}{\Delta_{i-1}},$$

where $\Delta_0 = 1$.

**Example 7.4.** Consider $Q(x_1, x_2, x_3) = x_1^2 - 4x_1x_2 + 2x_1x_3 + 6x_2^2 + 2x_3^2$.

$$A = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 6 & 0 \\ 1 & 0 & 2 \end{pmatrix},$$

$$\Delta_1 = 1, \;\; \Delta_2 = \begin{vmatrix} 1 & -2 \\ -2 & 6 \end{vmatrix} = 6 - 4 = 2,$$

$$\Delta_3 = \begin{vmatrix} 1 & -2 & 1 \\ -2 & 6 & 0 \\ 1 & 0 & 2 \end{vmatrix} = 12 - 6 - 8 = -2,$$

$$\lambda_1 = 1, \; \lambda_2 = \frac{2}{1} = 2, \;\; \lambda_3 = \frac{-2}{2} = -1,$$

in accordance with the canonical form obtained in example 7.3.

Note that every positive definite quadratic form $Q(x_1, x_2, \ldots, x_n)$ is turned into a negative definite quadratic form when multiplied by $-1$ (indeed, if $Q(x_1, x_2, \ldots, x_n) > 0$ for $(x_1, x_2, \ldots, x_n) \neq 0$ then $-Q(x_1, x_2, \ldots, x_n) < 0$ for $(x_1, x_2, \ldots, x_n) \neq 0$). Considering that the matrix of a quadratic form is also multiplied by $-1$,

$$\Delta_n = \begin{vmatrix} -a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & -a_{22} & \ldots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \ldots & -a_{nn} \end{vmatrix} = (-1)^{n^2} \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix},$$

we get that $\Delta_n(-Q) = (-1)^{n^2}\Delta_n(Q) = (-1)^n\Delta_n(Q)$ because $n^2$ is an even number if and only if $n$ also is. We thus get the following re-formulation of theorem 26 for negative-definite quadratic forms.

**Theorem 28.** *Let $Q(x) = x^TAx$ be a quadratic form in $n$ variables where $A = A^T$. The form $Q(x)$ is negative definite if and only if the upper left corner $k \times k$ minors of $A$ form an alternating series starting with a negative number, i.e., $\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \ldots$ etc.*

Note that Sylvester's criterion is a useful tool for the analysis of quadratic forms, but it is not always applicable because some of the upper left corner minors can be equal to zero. In the latter case one should use least squares completion.

**Example 7.5.** Consider $Q(x_1, x_2) = x_1x_2$. The Sylvester's criterion cannot be used as $\Delta_1 = 0$. However, the linear transformation $x_1 = y_1 - y_2$ and $x_2 = y_1 + y_2$ gives $Q(y_1, y_2) = (y_1 - y_2)(y_1 + y_2) = y_1^2 - y_2^2$, and we can safely conclude that the quadratic form $Q(x_1, x_2)$ is not definite.

## 7.5 Exercises

PROBLEM 7.1. Use the completion of full squares to find the canonical representation of the following quadratic forms.

(a) $-4x_1^2 + 8x_1x_2 + 10x_1x_3 - 2x_2^2 - 10x_2x_3 - 5x_3^2$

(b) $-x_1^2 - 6x_1x_2 - 2x_1x_3 + 4x_2^2 + 8x_2x_3 + 2x_3^2$

(c) $-5x_1^2 - 4x_1x_2 + 10x_1x_3 - 2x_2^2 + 10x_2x_3 - 5x_3^2$

(d) $x_1^2 + 4x_1x_2 + 10x_1x_3 + 2x_2^2 + 8x_2x_3 + 5x_3^2$

(e) $2x_1^2 - 4x_1x_2 + 8x_1x_3 - 4x_1x_4 - x_2^2 + 4x_2x_3 + 4x_2x_4 - x_3^2 - 2x_3x_4$

(f) $-x_1^2 - 6x_1x_2 - 6x_1x_3 - 8x_1x_4 + 5x_2^2 + 6x_2x_4 + 3x_3^2 + 6x_3x_4 + 4x_4^2$

(g) $x_1^2 - 4x_1x_2 - 2x_1x_3 + 5x_2^2 - 4x_2x_3 + 4x_2x_4 - x_3^2 - 4x_3x_4 + 4x_4^2$

(h) $2x_1^2 + 6x_1x_2 - 2x_1x_3 - 8x_1x_4 - 3x_2^2 - 6x_2x_3 + x_3^2 + 8x_3x_4 + 5x_4^2$

PROBLEM 7.2. Use Sylvester's criterion to find signatures of the following quadratic forms

(a) $5x_1^2 - 6x_1x_2 - 6x_1x_3 + 3x_2^2 + 6x_2x_3 + 3x_3^2$

(b) $3x_1^2 - 6x_1x_2 + 8x_1x_3 + x_2^2 - 6x_2x_3 + 4x_3^2$

(c) $-3x_1^2 - 6x_1x_2 + 6x_1x_3 + 6x_2x_3 + x_3^2$

(d) $-x_1^2 + 8x_1x_2 - 8x_1x_3 - 4x_2^2 + 8x_2x_3 - x_3^2$

(e) $-x_1^2 - 2x_1x_2 - 8x_1x_3 - 6x_1x_4 - 2x_2^2 + 2x_2x_3 + 4x_3^2 + 6x_3x_4 - 3x_4^2$

(f) $-2x_1^2 - 4x_1x_2 - 4x_1x_3 + 2x_2^2 + 4x_2x_3 - 8x_2x_4 + x_3^2 - 10x_3x_4 + x_4^2$

(g) $-x_1^2 - 2x_1x_3 - 4x_2^2 - 8x_2x_4 - 3x_3^2 - 4x_3x_4 - 9x_4^2$

(h) $-4x_2^2 + 16x_2x_4 - 3x_3^2 - 20x_4^2$

# 8 Euclidean vector spaces

The notion of a bilinear form occupies central place in the theory of Eucledian vector spaces.

**Definition 26.** A *bilinear form* on a real vector space $V$ is the mapping $b : V \times V \to \mathbb{R}$ such that

1. $\forall x_1, x_2, y \in V \;\; b(x_1 + x_2, y) = b(x_1, y) + b(x_2, y)$,
2. $\forall x, y_1, y_2 \in V \;\; b(x, y_1 + y_2) = b(x, y_1) + b(x, y_2)$,

3. $\forall x, y \in V, \lambda \in \mathbb{R}\;\; b(\lambda x, y) = b(x, \lambda y) = \lambda b(x, y)$.

In other words, $b(x, y)$ is a linear function of both its arguments. The bilinear form $b(x, y)$ is called *symmetric* if

4. $\forall x, y \in V\;\; b(x, y) = b(y, x)$.

The bilinear form $b(x, y)$ is called *positive definite* if

5. $\forall x \in V\;\; b(x, x) \geq 0$ and $b(x, x) = 0$ only if $x = 0$.

Similarly to what was discussed in the previous section, each bilinear form admits the coordinate representation, in which it is expressed by using matrix products:

$$b(x, y) = x^T B y,$$

where $x, y \in \mathbb{R}^n$, and $B$ is a n×n matrix, which is called the *matrix of a bilinear form.* Obviously, the bilinear form is symmetric if and only if its matrix is symmetric. The change of base for bilinear forms occurs similarly to what we saw for quadratic forms. If $x = Cx'$ then $b(x, y) = x^T B y = (x')^T C^T B C x'$, i.e., the matrix of a bilinear form also transforms by the rule $B \mapsto C^T B C$.

Each symmetric bilinear form $b(x, y)$ defines a quadratic form with the same matrix by the identity $Q(x) = b(x, x)$. Conversely, if $Q(x) = x^T A x$ is a quadratic form then we can define a bilinear form by using the same matrix, i.e., $b(x, y) = x^T A y$. This can be done even without matrices because since $b(x + y, x + y) = b(x, x) + b(x, y) + b(y, x) + b(y, y) = Q(x) + Q(y) + 2b(x, y) = Q(x + y)$ and

$$b(x, y) = \frac{1}{2}\left(Q(x + y) - Q(x) - Q(y)\right).$$

Thus, each quadratic form defines a symmetric bilinear form and, vice versa, each symmetric bilinear form defines a quadratic form.

## 8.1 Linear spaces with dot product

A symmetric, positive definite bilinear form $b(x, y)$ on a real vector space $V$ is called *dot product* (also called *scalar product*). By theorem 24, there is a base in $V$ such that the corresponding to $b(x, y)$ quadratic form $Q(x)$ (one with the same symmetric matrix) has the form $x_1^2 + x_2^2 + \cdots + x_n^2$, i.e., in some base of $V$, the matrix of $Q(x)$ is the identity matrix. In the same base the matrix of $b(x, y)$ must also be the identity matrix, i.e., the bilinear form is expressed as[14]

$$b(x, y) = x^T y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

A linear vector space, along with a dot product defined on it, is called a *Euclidean vector space.*

The scalar product in the Euclidean vector space can be considered as a "source of geometry". Indeed, the colloquial dot product of vectors can be used to define lengths (by setting $l(x) =$

[14] Note that the product of a 1×n (row) and n×1 (column) matrices gives a 1×1 matrix, i.e. a number.

$\sqrt{Q(x)} = \sqrt{b(x,x)})$ and angles (by setting the angle between vectors $x$ and $y$ be equal to $\alpha(x,y) = \cos^{-1}\frac{b(x,y)}{l(x)l(y)})$.

In what follows we will discuss well-known Euclidean vector spaces with a dot product which is so "standard" that we omit the name of the function, i.e., instead of $b(x,y)$ we write $(x,y)$.

**Example 8.1.** The $n$-dimensional real vector space $\mathbb{R}^n$ is a Euclidean vector space with respect to the dot product defined by

$$(x,y) = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

**Example 8.2.** The vector space $C[a,b]$ of continuous real functions defined on $[a,b]$ is a Euclidean vector space with respect to the dot product defined by

$$(f,g) = \int_a^b f(x)\cdot g(x)dx.$$

Indeed, it is linear, symmetric, and positive definite because

$$(f,f) = \int_a^b f^2(x)dx \geq 0,$$

and it could be equal to zero only if $f(x)$ differs from zero on a zero-measure set, which implies that $f(x) = 0$ for a continuous function.

**Example 8.3.** Consider the space $V$ of discrete random numbers (a random number consists of a finite collection of values, each endowed with a number called probability, such that all probabilities are positive and sum up to 1). It is a linear vector space with respect to the addition of random numbers and multiplication by scalars. There is a linear operator of expected value acting on $V$. It is used to introduce variance and covariance as follows:

$$\operatorname{cov}(X,Y) = E(X\cdot Y) - E(X)\cdot E(Y),$$
$$\operatorname{var}(X) = \operatorname{cov}(X,X) = E(X^2) - E^2(X).$$

Here, $\operatorname{cov}(X,Y)$ brings Euclidean structure to the vector space $V$, thus establishing an association between vectors and random numbers, one in which the length of a random number (as of a vector) is $\sqrt{\operatorname{cov}(X,X)} = \sqrt{\operatorname{var}(X)} = \sigma(X)$, i.e., the length is somewhat similar to standard deviation, while the cosine of the angle between two random numbers is their correlation. Here, independent random numbers, which necessarily should have zero covariance, turn out to be orthogonal, i.e., the angle between them is 90°.

## 8.2 Gram-Schmidt orthogonalization

**Definition 27.** Vectors $x$ and $y$ of a Euclidean vector space are called *orthogonal* if their dot product is equal to zero, i.e. if $(x,y) = 0$. A base of $V$ consisting of vectors $e_1, e_2, \ldots, e_n$ such that $(e_i, e_j) = 0$ for $i \neq j$ is called an *orthogonal base*. An orthogonal base is called *orthonormal* if, additionally, $(e_i, e_i) = 1$ for all $i$.

An orthogonal base consists of vectors that are perpendicular to each other. The vectors forming an orthonormal base are, in addition to this, each of length one. Obviously, each orthogonal base can be transformed to the orthonormal base by dividing its vector by their lengths. It is very convenient to deal with an orthonormal base because the components of the vector $x = x_1e_1 + x_2e_2 + \cdots + x_ne_n$ in this base can be conveniently calculated by using the dot product:

$$(x, e_i) = (x_1e_1 + x_2e_2 + \cdots + x_ne_n, e_i) = 0 + \cdots + x_i(e_i, e_i) + \cdots + 0 = x_i.$$

If $e_1, e_2, \ldots, e_n$ is a (not necessarily orthogonal) base in $V$ then the matrix $G = (g_{ij})$, where $g_{ij} = (e_i, e_j)$ is called the *Gram matrix* of this base. Each Gram matrix is symmetric and positive definite because so is the dot product. The Gram matrix of an orthogonal base is diagonal; the Gram matrix of an orthonormal base is the identity matrix. That is, if $e_1, e_2, \ldots, e_n$ is an orthonormal base then $(e_i, e_j) = \delta_{ij}$, where $\delta_{ij}$ is so called Kronecker's symbol such that $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise.

Let $e_1, e_2, \ldots, e_n$ be a base in $V$ which is not orthogonal. The procedure below, called Gram-Schmidt orthogonalization process, can be used to transform $(e_1, e_2, \ldots, e_n)$ to an orthonormal base $(f_1, f_2, \ldots, f_n)$ so that the flags of linear hulls generated by these two bases coincide, i.e. for all $i$

$$\begin{aligned} \mathcal{L}(e_1) &= \mathcal{L}(f_1), \\ \mathcal{L}(e_1, e_2) &= \mathcal{L}(f_1, f_2), \\ \mathcal{L}(e_1, \ldots, e_i) &= \mathcal{L}(f_1, \ldots, f_i) \end{aligned}$$

for all $i$. We will construct $(f_1, f_2, \ldots, f_n)$ in a series of steps. First, let $f_1 = e_1$. Assume that $f_1, f_2, \ldots, f_{k-1}$ have been constructed and we want to find $f_k$ such that $f_1, f_2, \ldots, f_k$ are pairwise orthogonal and $\mathcal{L}(f_1, \ldots, f_k) = \mathcal{L}(e_1, \ldots, e_k)$. We will search for $f_k$ in a form of linear combination of $e_k$ and already constructed vectors, i.e.,

$$f_k = e_k - \lambda_1 f_1 - \lambda_2 f_2 - \cdots - \lambda_{k-1} f_{k-1}.$$

We need that $(f_k, f_1) = (f_k, f_2) = \cdots = (f_k, f_{k-1}) = 0$. Then,

$$\begin{aligned} (e_k - \lambda_1 f_1 - \lambda_2 f_2 - \cdots - \lambda_{k-1} f_{k-1}, f_1) &= (e_k, f_1) - \lambda_1 (f_1, f_1) = 0, \\ (e_k - \lambda_1 f_1 - \lambda_2 f_2 - \cdots - \lambda_{k-1} f_{k-1}, f_2) &= (e_k, f_2) - \lambda_1 (f_2, f_2) = 0, \\ (e_k - \lambda_1 f_1 - \lambda_2 f_2 - \cdots - \lambda_{k-1} f_{k-1}, f_{k-1}) &= (e_k, f_{k-1}) - \lambda_1 (f_{k-1}, f_{k-1}) = 0. \end{aligned}$$

That is,

$$\lambda_i = \frac{(e_k, f_i)}{(f_i, f_i)}.$$

**Example 8.4.** Use Gram-Schmidt orthogonalization process to find an orthogonal base in the linear hull of the following system of vectors

$$e_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \ e_2 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \ e_3 = \begin{pmatrix} 1 \\ -4 \\ 0 \end{pmatrix}$$

**Solution.** As explained above, we first set $f_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Then, we look for $f_2$ in the following form

$$f_2 = e_2 - \lambda_1 f_1.$$

We have $(f_2, f_1) = (e_2, f_1) - \lambda_1(f_1, f_1) = 0$. Then, $(e_2, f_1) = 4$, $(f_1, f_1) = 3$, $\lambda_1 = \frac{4}{3}$, and $f_2 = e_2 - \frac{4}{3} f_1 = \begin{pmatrix} \frac{5}{3} \\ -\frac{1}{3} \\ -\frac{4}{3} \end{pmatrix}$. One may want to stretch $f_2$ three times to obtain integer components, $f_2 = \begin{pmatrix} 5 \\ -1 \\ -4 \end{pmatrix}$.

Next, we look for $f_3$ as $f_3 = e_3 - \lambda_1 f_1 - \lambda_2 f_2$. Then, $(f_3, f_1) = (e_3, f_1) - \lambda_1(f_1, f_1) - \lambda_2(f_2, f_1) = (e_3, f_1) - \lambda_1(f_1, f_1) = 0$. We have $(e_3, f_1) = -3$, $(f_1, f_1) = 3$. Therefore, $\lambda_1 = -1$. Similarly, $(f_3, f_2) = (e_3, f_2) - \lambda_1(f_1, f_2) - \lambda_2(f_2, f_2) = (e_3, f_2) - \lambda_2(f_2, f_2) = 0$. Here, $(e_3, f_2) = 9$, $(f_2, f_2) = 42$, and $\lambda_2 = 9/42 = 3/14$. Then, we have $f_3 = e_3 + f_1 - \frac{3}{14} f_2 = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} - \begin{pmatrix} \frac{15}{14} \\ -\frac{3}{14} \\ -\frac{12}{14} \end{pmatrix} = \begin{pmatrix} \frac{13}{14} \\ -\frac{39}{14} \\ \frac{26}{14} \end{pmatrix}$, which gives $f_3 = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}$ after stretching $14/13$ times.

## 8.3 Projection of a vector on a linear subspace

In many parts of mathematics appears the problem of the optimal solution or, more naïvely, the problem of approximation or finding the "closest point". For instance, the linear regression in Statistics deals with fitting a line through the points which are not on the same line; still, there exists the "optimal" line, which the the closest one among all lines in the plane with respect to the given set of measurements.

Below we provide a general framework in the context of Euclidean space. Suppose you have a vector $x$ in the Euclidean space $V$ and a vector subspace $U \subseteq V$ generated by a set of vectors, $U = \mathcal{L}(a_1, a_2, \dots, a_k)$. If $x \notin U$ then it cannot be represented as a linear combination $\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_k a_k$. However, one can ask when the length of the difference $x - \lambda_1 a_1 - \lambda_2 a_2 - \dots - \lambda_k a_k$ is minimal. This can happen if and only if the vector $x - \lambda_1 a_1 - \lambda_2 a_2 - \dots - \lambda_k a_k$ is orthogonal[15] to the subspace $U$. Denote the difference vector by $n$, i.e.,

$$n = x - \lambda_1 a_1 - \lambda_2 a_2 - \dots - \lambda_k a_k.$$

It will be called the *normal* vector[16]. The normal vector must be orthogonal to each of the

[15] Can you explain, why? Note that your colloquial reasoning such as "the shortest path is the perpendicular" in general fails to work.

[16] Here "normal" is the opposite to "tangent".

vectors in $U$, including $a_1, a_2, \dots, a_k$. Then, we have

$$\begin{cases} (n, a_1) = (x, a_1) - \lambda_1(a_1, a_1) - \lambda_2(a_2, a_1) - \dots - \lambda_n(a_k, a_1) = 0 \\ (n, a_2) = (x, a_2) - \lambda_1(a_1, a_2) - \lambda_2(a_2, a_2) - \dots - \lambda_n(a_k, a_2) = 0 \\ \dots \\ (n, a_k) = (x, a_k) - \lambda_1(a_1, a_k) - \lambda_2(a_2, a_k) - \dots - \lambda_n(a_k, a_k) = 0 \end{cases}.$$

Therefore,

$$\begin{cases} \lambda_1(a_1, a_1) + \lambda_2(a_2, a_1) + \dots + \lambda_n(a_k, a_1) = (x, a_1) \\ \lambda_1(a_1, a_2) + \lambda_2(a_2, a_2) + \dots + \lambda_n(a_k, a_2) = (x, a_2) \\ \dots \\ \lambda_1(a_1, a_k) + \lambda_2(a_2, a_k) + \dots + \lambda_n(a_k, a_k) = (x, a_k) \end{cases},$$

that is, the set of coefficients $\lambda_1, \lambda_2, \dots, \lambda_k$ is obtained from solving the system of linear equations whose matrix is the Gram matrix of $a_1, a_2, \dots, a_k$ and the column of free terms consists of dot products of $x$ with $a_1, a_2, \dots, a_k$.

**Example 8.5.** For the vectors $x$, $e_1$, and $e_2$ below, find the projection of the vector $x$ onto the linear hull of $e_1$ and $e_2$. Evaluate the distance between the endpoint of $x$ and the linear hull.

$$x = \begin{pmatrix} -1 \\ -7 \\ -1 \end{pmatrix}, \ e_1 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \ e_2 = \begin{pmatrix} -2 \\ 0 \\ -6 \end{pmatrix}$$

**Solution.** Compute the Gram matrix: $(e_1, e_1) = 9$, $(e_1, e_2) = -16$, $(e_2, e_2) = 40$, $(x, e_1) = -11$, $(x, e_2) = 8$. Then,

$$\left(\begin{array}{rr|r} 9 & -16 & -11 \\ -16 & 40 & 8 \end{array}\right) \to \left(\begin{array}{rr|r} 2 & 8 & -14 \\ -16 & 40 & 8 \end{array}\right) \to \left(\begin{array}{rr|r} 1 & 4 & -7 \\ -16 & 40 & 8 \end{array}\right) \to$$

$$\left(\begin{array}{rr|r} 1 & 4 & -7 \\ 0 & 104 & -104 \end{array}\right) \to \left(\begin{array}{rr|r} 1 & 4 & -7 \\ 0 & 1 & -1 \end{array}\right) \to \left(\begin{array}{rr|r} 1 & 0 & -3 \\ 0 & 1 & -1 \end{array}\right).$$

Therefore, $\lambda_1 = -3$, $\lambda_2 = -1$, and

$$\lambda_1 e_1 + \lambda_2 e_2 = -3 \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} - \begin{pmatrix} -2 \\ 0 \\ -6 \end{pmatrix} = \begin{pmatrix} -4 \\ -3 \\ 0 \end{pmatrix}.$$

$$n = \begin{pmatrix} -1 \\ -7 \\ -1 \end{pmatrix} - \begin{pmatrix} -4 \\ -3 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix},$$

and $d^2 = 3^2 + 4^2 + 1^2 = 26$.

**Example 8.6.** For the vectors $x$, $e_1$, $e_2$, and $e_3$ below, find the projection of the vector $x$ onto the linear hull of $e_1$, $e_2$, and $e_3$. Evaluate the distance between the endpoint of $x$ and the linear hull.

$$x = \begin{pmatrix} 1 \\ 6 \\ -1 \\ 1 \end{pmatrix}, \ e_1 = \begin{pmatrix} -2 \\ 2 \\ -1 \\ -3 \end{pmatrix}, \ e_2 = \begin{pmatrix} -4 \\ 4 \\ -3 \\ -5 \end{pmatrix}, \ e_3 = \begin{pmatrix} -3 \\ 4 \\ -3 \\ -5 \end{pmatrix}.$$

**Solution.** Compute the Gram matrix: $(e_1, e_1) = 18$, $(e_1, e_2) = 34$, $(e_1, e_3) = 32$, $(e_2, e_2) = 66$, $(e_2, e_3) = 62$,$(e_3, e_3) = 59$, $(x, e_1) = 8$, $(x, e_2) = 18$, $(x, e_3) = 19$. Then,

$$\left(\begin{array}{ccc|c} 18 & 34 & 32 & 8 \\ 34 & 66 & 62 & 18 \\ 32 & 62 & 59 & 19 \end{array}\right) \to \left(\begin{array}{ccc|c} 9 & 17 & 16 & 4 \\ 17 & 33 & 31 & 9 \\ 32 & 62 & 59 & 19 \end{array}\right) \xrightarrow{2[1]-[2]} \dots$$

$$\lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3 = \begin{pmatrix} 1 \\ 2 \\ -3 \\ -1 \end{pmatrix}, \; n = \begin{pmatrix} 0 \\ 4 \\ 2 \\ 2 \end{pmatrix}, \text{ and}$$

$$d^2 = 24.$$

**Example 8.7.** For the function $y = x^3$, find its projection onto the linear space generated by 1, $x$ and $x^2$ with respect to Euclidean structure defined by

$$(f, g) = \int_0^1 f(x)g(x)dx$$

.

**Solution.** $(1, 1) = 1$, $(1, x) = \int_0^1 x dx = 1/2$, $(x, x) = (1, x^2) = \int_0^1 x^2 dx = 1/3$, $(1, x^3) = (x, x^2) = \int_0^1 x^3 dx = 1/4$, $(x^2, x^2) = (x, x^3) = \int_0^1 x^4 dx = 1/5$, $(x^2, x^3) = \int_0^1 x^5 dx = 1/6$.

$$\left(\begin{array}{ccc|c} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \end{array}\right) \to \left(\begin{array}{ccc|c} 12 & 6 & 4 & 3 \\ 30 & 20 & 15 & 12 \\ 40 & 30 & 24 & 20 \end{array}\right) \to \left(\begin{array}{ccc|c} 12 & 6 & 4 & 3 \\ 30 & 20 & 15 & 10 \\ 1 & 1 & .9 & .8 \end{array}\right) \to \dots$$

$$\lambda_1 = .05, \lambda_1 = -.6, \lambda_2 = 1.5,$$

i.e., the projection of $y = x^3$ onto the linear space generated by 1, $x$ and $x^2$ is $y = .05 - .6x + 1.5x^2$. Spend some time with your graphic calculator to draw graphs of these two functions.

## 8.4 Orthogonal complement

Euclidean structure on a linear vector space gives rise to the notion of orthogonal subspaces, which in our colloquial understanding corresponds to perpendicular planes. Let $V$ be a Euclidean vector space and let $U \subseteq V$ be its subspace.

**Definition 28.** The orthogonal complement to $U$ in $V$ is a subset

$$U^\perp = \{x \in V | \forall y \in U : (x, y) = 0\},$$

i.e., $U^\perp$ consists of all vectors of $V$ that are orthogonal to every vector in $U$.

For instance, if $U$ is the line defined by $x = y$ in $\mathbb{R}^2$ then $U^\perp$ is the line $x + y = 0$, which is perpendicular to the original line. If $U$ is the line defined by $x = y = z$ in $\mathbb{R}^3$ then $U^\perp$ is the plane that passes through the origin at 90° to the original line.

It follows immediately from this definition that $U^\perp$ is a correctly defined subspace of $V$ because if $x_1$ and $x_2$ are such that $\forall y \in U : (x_1, y) = (x_2, y) = 0$ then $x_1 \pm x_2$ also have the same property due to linear nature of the dot product.

**Theorem 29.**

$$\dim(U) + \dim(U^\perp) = \dim(V)$$

*Proof.* One can choose an arbitrary base $e_1, \ldots, e_k$, where $k = \dim(U)$, in $U$ and complement it by vectors $e_{k+1}, \ldots, e_n$ so that $e_1, \ldots, e_n$ is a base of $V$. Applying the Gram-Schmidt orthogonalization procedure, we get an orthogonal base $e'_1, \ldots, e'_n$ of $V$ such that $\mathcal{L}(e'_1, \ldots, e'_k) = \mathcal{L}(e_1, \ldots, e_k) = U$. Let $x$ be a vector that is orthogonal to all vectors in $U$ and let $x = x_1e'_1 + x_2e'_2 + \cdots + x_ne'_n$ be its decomposition over the base of $V$. Then, $(x, e'_i) = x_i = 0$ for all $1 \le i \le k$. Therefore, $x = x_{k+1}e'_{k+1} + \cdots + x_ne'_n$, i.e., $e'_{k+1}, \ldots, e'_n$ is the base of $U^\perp$. Therefore, $\dim(U^\perp) = n - k = n - \dim(U)$. □

**Corollary 29.1.** $\left(U^\perp\right)^\perp = U$

*Proof.* It is obvious that $U$ is a subset of $\left(U^\perp\right)^\perp$. By the previous theorem, we get that $\dim\left(\left(U^\perp\right)^\perp\right) = \dim(V) - \dim(U^\perp) = \dim(V) - (\dim(V) - \dim(U)) = \dim(U)$, i.e., $U = \left(U^\perp\right)^\perp$. □

**Theorem 30.** *If $U_1 \subseteq U_2$ then $U_2^\perp \subseteq U_1^\perp$.*

*Proof.* If $x \in U_2^\perp$ then $(x, y) = 0$ for all $y \in U_2$. Since every vector of $U_1$ also belongs to $U_2$, we have $x \in U_1^\perp$, i.e., $U_2^\perp \subseteq U_1^\perp$. □

**Example 8.8.** Find the orthogonal complement to the linear vector space generated by $e_1$ and $e_2$ in $\mathbb{R}^3$, where $e_1 = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$ and $e_2 = \begin{pmatrix} 5 \\ -1 \\ -10 \end{pmatrix}$.

**Solution.** If $x = (x_1, x_2, x_3)$ be a vector in the orthogonal complement, then

$$\begin{cases} x_1 + 3x_2 - 2x_3 = 0 \\ 5x_1 - x_2 - 10x_3 = 0 \end{cases},$$

as $x$ must be orthogonal to both $e_1$ and $e_2$. We get

$$\begin{pmatrix} 1 & 3 & -2 \\ 5 & -1 & -10 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & -2 \\ 0 & -16 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & -2 \\ 0 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \end{pmatrix}.$$

The free variable is $x_3$. By setting $x_3 = 1$, we get $x_1 = 2$ and $x_2 = 0$. Hence, the vectors that belong to the orthogonal complement must be proportionate to $x$, i.e., the orthogonal complement is the line generated by $x = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$. □

## 8.5 Intersection and sum of vector spaces

Given two linear subspaces, $U_1$ and $U_2$, of a vector space $V$, one may want to consider their union and intersection, $U_1 \cup U_2$ and $U_1 \cap U_2$. While the latter is a correctly defined linear subspace (if $x, y \in U_1 \cap U_2$ then $x, y \in U_1$ and $x, y \in U_2$, and thus $x + y \in U_1$ and $x + y \in U_2$, i.e., $x + y \in U_1 \cap U_2$), the former is not a linear subspace because $x \in U_1$ and $y \in U_2$ does not imply that $x + y$ belongs to either of the two subspaces.

However, one can define the sum $U_1 + U_2$ of two linear subspaces to be not precisely the union, but the linear hull of $U_1 \cup U_2$. This way we "force" it to be a linear subspace. Following this definition, $U_1 + U_2$ is a bigger set than $U_1 \cup U_2$, but it is the smallest of all linear subspaces of $V$ that contain $U_1 \cup U_2$.

If $U_1$ and $U_2$ are defined by their bases, it is easy to find the base of $U_1 + U_2$. Indeed, if $U_1 = \mathcal{L}(e_1, e_2, \ldots, e_k)$ and $U_2 = \mathcal{L}(f_1, f_2, \ldots, f_l)$ then $U_1 + U_2$ is generated by $e_1, e_2, \ldots, e_k, f_1, f_2, \ldots, f_l$. If they are not linearly independent, we already know how to select a base (section 2.4).

**Example 8.9.** Find the sum of the vector spaces $U_1 = \mathcal{L}(e_1, e_2)$ and $U_2 = \mathcal{L}(f_1, f_2)$, where

$$e_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \ e_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \ f_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ -1 \end{pmatrix}, \text{ and } f_2 = \begin{pmatrix} 2 \\ -1 \\ 3 \\ 2 \end{pmatrix}.$$

**Solution.** The problem is equivalent to building a base from a set vectors, i.e., we have to examine whether or not $e_1, e_2, f_1, f_2$ are linearly independent and, if not, choose a base of their linear hull.

$$\begin{pmatrix} 2 & 0 & 1 & 2 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 2 & 3 \\ 0 & 1 & -1 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 2 & -1 & 4 \\ 0 & 2 & 1 & 4 \\ 0 & 1 & -1 & 2 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix} \to \begin{pmatrix} \underline{1} & -1 & 1 & -1 \\ 0 & \underline{1} & -1 & 2 \\ 0 & 0 & \underline{1} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

In this case, the vectors $e_1, e_2, f_1, f_2$ are linearly dependent, but $e_1, e_2, f_1$ form a base, so the sum of two 2-dimensional subspaces $U_1$ and $U_2$ is a three dimensional subspace spanned on $e_1$,$e_2$, and $f_1$. It means that $U_1$ and $U_2$ are "planes" in 4D situated so that they have a common "line" and $U_1 + U_2$ is three-dimensional.

**Theorem 31.**

$$\dim(U_1 + U_2) = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2).$$

*Proof.* Chose a base in $U_1 \cap U_2$. One can add vectors to this base to make it the base of $U_1$ since $U_1 \cap U_2 \subseteq U_1$. Similarly, one can add vectors to the base of $U_1 \cap U_2$ to make it the base of $U_2$ as $U_1 \cap U_2 \subseteq U_2$, too. If we now combine the two bases, but without counting the base vectors of $U_1 \cap U_2$ twice, we get the base of $U_1 + U_2$, which implies that $\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = \dim(U_1) + \dim(U_2)$. □

In the particular case when $\dim(U_1 \cap U_2) = 0$, which is possible only when $U_1 \cap U_2 = \{0\}$, i.e., the zero vector is the only intersection of $U_1$ and $U_2$, the sum of vector spaces is called a *direct sum* and is denoted by $U_1 \oplus U_2$. For the direct sum of vector spaces, the dimension formula becomes

$$\dim(U_1 \oplus U_2) = \dim(U_1) + \dim(U_2).$$

Computations are not that straightforward for the intersection. Indeed, if $U_1$ and $U_2$ are the same as in example 8.9, the base vectors of $U_1 \cap U_2$ could be neither of $e_1$,$e_2$, $f_1$, or $f_2$. However, it can be found very easily by using orthogonal complements and corollary 8.4.

**Theorem 32.**

$$(U_1 + U_2)^{\perp} = U_1^{\perp} \cap U_2^{\perp}.$$

*Proof.* Since $U_1 \subseteq U_1 \cup U_2 \subseteq U_1 + U_2$, we have $(U_1+U_2)^{\perp} \subseteq U_1^{\perp}$. Similarly, we have $(U_1+U_2)^{\perp} \subseteq U_2^{\perp}$ and, therefore, $(U_1 + U_2)^{\perp} \subseteq U_1^{\perp} \cap U_2^{\perp}$. On the other hand, if $x \in U_1^{\perp} \cap U_2^{\perp}$ then $(x, y_1) = 0$ for every $y_1 \in U_1$ and also $(x, y_2) = 0$ for every $y_2 \in U_2$. Therefore, $U_1^{\perp} \cap U_2^{\perp} \subseteq (U_1 + U_2)^{\perp}$, i.e., it is actually an equality. $\square$

**Theorem 33.**

$$(U_1 \cap U_2)^{\perp} = U_1^{\perp} + U_2^{\perp}.$$

*Proof.* Similar to the proof of theorem 8.5. $\square$

**Example 8.10.** Find the intersection of vector spaces $U_1 = \mathcal{L}(e_1, e_2)$ and $U_2 = \mathcal{L}(f_1, f_2)$, where $e_1$, $e_2$, $f_1$, and $f_2$ are the same as in example 8.9.

**Solution.** First, let us find $U_1^{\perp}$ and $U_2^{\perp}$.

$$\begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & -1 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 2 & 0 & 2 & 1 \\ 0 & -1 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 & \frac{1}{2} \\ 0 & 1 & -1 & -1 \end{pmatrix}$$

The base of $U_1^{\perp}$ is $a_1 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$ and $a_2 = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 2 \end{pmatrix}$. Similarly,

$$\begin{pmatrix} 1 & 1 & 2 & -1 \\ 2 & -1 & 3 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 2 & -1 \\ 0 & -3 & -1 & 4 \end{pmatrix} \to$$

$$\begin{pmatrix} 3 & 3 & 6 & -3 \\ 0 & -3 & -1 & 4 \end{pmatrix} \to \begin{pmatrix} 3 & 0 & 5 & 1 \\ 0 & 3 & 1 & -4 \end{pmatrix}$$

The base of $U_2^{\perp}$ is $b_1 = \begin{pmatrix} -5 \\ -1 \\ 3 \\ 0 \end{pmatrix}$ and $b_2 = \begin{pmatrix} -1 \\ 4 \\ 0 \\ 3 \end{pmatrix}$. Now we find the base of their sum and, simultaneously, the base of $\left(U_1^{\perp} + U_2^{\perp}\right)^{\perp}$.

$$\begin{pmatrix} -1 & 1 & 1 & 0 \\ -1 & 2 & 0 & 2 \\ -5 & -1 & 3 & 0 \\ -1 & 4 & 0 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & -1 & 0 \\ 0 & 1 & -1 & 2 \\ 0 & -6 & -2 & 0 \\ 0 & 3 & -1 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & -1 & 0 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & -8 & 12 \\ 0 & 0 & 2 & -3 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 2 & -2 & -2 & 0 \\ 0 & 2 & -2 & 4 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 2 & -2 & 0 & -3 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 2 & -3 \end{pmatrix} \to \begin{pmatrix} 2 & 0 & 0 & -2 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 2 & -3 \end{pmatrix}.$$

Thus, the base of $\left(U_1^\perp + U_2^\perp\right)^\perp = U_1 \cap U_2$ is $x = \begin{pmatrix} 2 \\ -1 \\ 3 \\ 2 \end{pmatrix}$, which is the same as the vector $f_2$, and which turns out to be the direction vector of the line that is common to both $U_1$ and $U_2$. □

## 8.6 Exercises

PROBLEM 8.1. Use Gram-Schmidt orthogonalization process to find an orthogonal base in the linear hull of the following system of vectors

(a) $e_1 = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}, e_2 = \begin{pmatrix} -1 \\ 1 \\ 4 \end{pmatrix}, e_3 = \begin{pmatrix} -4 \\ 2 \\ 1 \end{pmatrix}$

(b) $e_1 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}, e_2 = \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}, e_3 = \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}$

(c) $e_1 = \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix}, e_2 = \begin{pmatrix} -1 \\ 1 \\ -3 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix}$

(d) $e_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} -4 \\ 0 \\ 1 \end{pmatrix}, e_3 = \begin{pmatrix} -3 \\ -2 \\ -1 \end{pmatrix}$

(e) $e_1 = \begin{pmatrix} 2 \\ 2 \\ 3 \\ 1 \end{pmatrix}, e_2 = \begin{pmatrix} 2 \\ 3 \\ 3 \\ 2 \end{pmatrix}, e_3 = \begin{pmatrix} 1 \\ -2 \\ -2 \\ -2 \end{pmatrix}$

(f) $e_1 = \begin{pmatrix} 3 \\ 4 \\ 1 \\ 2 \end{pmatrix}, e_2 = \begin{pmatrix} 4 \\ 1 \\ 3 \\ 3 \end{pmatrix}, e_3 = \begin{pmatrix} 4 \\ 2 \\ 3 \\ 1 \end{pmatrix}$

(g) $e_1 = \begin{pmatrix} 3 \\ 1 \\ 3 \\ 1 \\ 1 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ -1 \\ -4 \\ 1 \\ -3 \end{pmatrix}, e_3 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 0 \\ 3 \end{pmatrix}, e_4 = \begin{pmatrix} 2 \\ -4 \\ 3 \\ -2 \\ -4 \end{pmatrix}$

(h) $e_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 3 \\ 3 \end{pmatrix}, e_2 = \begin{pmatrix} 4 \\ -4 \\ -2 \\ 2 \\ 4 \end{pmatrix}, e_3 = \begin{pmatrix} 3 \\ -2 \\ -3 \\ 1 \\ 2 \end{pmatrix}, e_4 = \begin{pmatrix} -2 \\ -3 \\ -2 \\ 1 \\ 0 \end{pmatrix}$

PROBLEM 8.2. Find the Gram matrix of the following vectors

(a) $e_1 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, e_2 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, e_3 = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$

(b) $e_1 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}, e_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, e_3 = \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix}$

(c) $e_1 = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, e_2 = \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

(d) $e_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, e_2 = \begin{pmatrix} 3 \\ -2 \\ 2 \end{pmatrix}, e_3 = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}$

(e) $e_1 = \begin{pmatrix} 3 \\ 2 \\ 2 \\ 2 \end{pmatrix}, e_2 = \begin{pmatrix} -2 \\ 2 \\ -2 \\ 1 \end{pmatrix}, e_3 = \begin{pmatrix} 1 \\ -1 \\ -4 \\ -1 \end{pmatrix}, e_4 = \begin{pmatrix} -3 \\ -1 \\ -3 \\ 3 \end{pmatrix}$

(f) $e_1 = \begin{pmatrix} 1 \\ 2 \\ 4 \\ 3 \end{pmatrix}, e_2 = \begin{pmatrix} -1 \\ 3 \\ -2 \\ -1 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}, e_4 = \begin{pmatrix} 1 \\ 2 \\ -3 \\ 3 \end{pmatrix}$

(g) $e_1 = \begin{pmatrix} 0 \\ 1 \\ 3 \\ 2 \end{pmatrix}, e_2 = \begin{pmatrix} 3 \\ -1 \\ 3 \\ -3 \end{pmatrix}, e_3 = \begin{pmatrix} 2 \\ -1 \\ 2 \\ -1 \end{pmatrix}, e_4 = \begin{pmatrix} 1 \\ 1 \\ -4 \\ -4 \end{pmatrix}$

(h) $e_1 = \begin{pmatrix} 2 \\ 4 \\ 1 \\ 1 \end{pmatrix}$, $e_2 = \begin{pmatrix} -4 \\ 2 \\ -3 \\ 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 0 \\ -1 \\ -2 \\ -2 \end{pmatrix}$, $e_4 = \begin{pmatrix} 3 \\ 0 \\ 1 \\ -1 \end{pmatrix}$

PROBLEM 8.3. For the vectors $x, e_1, e_2, \ldots, e_k$ below, find the projection of the vector $x$ onto the linear hull of $e_1, e_2, \ldots, e_k$. Evaluate the distance between the endpoint of $x$ and the linear hull.

(a) $x = \begin{pmatrix} -1 \\ -6 \\ -2 \end{pmatrix}$, $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$

(b) $x = \begin{pmatrix} -2 \\ -7 \\ 4 \end{pmatrix}$, $e_1 = \begin{pmatrix} 4 \\ 2 \\ -4 \end{pmatrix}$, $e_2 = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$

(c) $x = \begin{pmatrix} 3 \\ 3 \\ -3 \end{pmatrix}$, $e_1 = \begin{pmatrix} -3 \\ 3 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}$

(d) $x = \begin{pmatrix} -4 \\ 2 \\ -3 \end{pmatrix}$, $e_1 = \begin{pmatrix} -1 \\ -5 \\ -4 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

(e) $x = \begin{pmatrix} 3 \\ 4 \\ 4 \\ 2 \end{pmatrix}$, $e_1 = \begin{pmatrix} 4 \\ 0 \\ 3 \\ -3 \end{pmatrix}$, $e_2 = \begin{pmatrix} -1 \\ 3 \\ -3 \\ 3 \end{pmatrix}$, $e_3 = \begin{pmatrix} -2 \\ -5 \\ 4 \\ -4 \end{pmatrix}$

(f) $x = \begin{pmatrix} 1 \\ -6 \\ -3 \\ -3 \end{pmatrix}$, $e_1 = \begin{pmatrix} -5 \\ -2 \\ -1 \\ 3 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ -3 \end{pmatrix}$, $e_3 = \begin{pmatrix} 3 \\ -2 \\ 0 \\ 4 \end{pmatrix}$

(g) $x = \begin{pmatrix} -4 \\ 2 \\ 0 \\ -2 \end{pmatrix}$, $e_1 = \begin{pmatrix} -4 \\ 1 \\ -5 \\ 4 \end{pmatrix}$, $e_2 = \begin{pmatrix} -5 \\ 4 \\ -5 \\ 5 \end{pmatrix}$, $e_3 = \begin{pmatrix} 1 \\ -4 \\ 0 \\ -1 \end{pmatrix}$

(h) $x = \begin{pmatrix} 5 \\ -4 \\ -3 \\ -4 \end{pmatrix}$, $e_1 = \begin{pmatrix} -2 \\ 4 \\ 3 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 2 \\ -1 \\ -3 \\ 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 1 \\ -5 \\ -3 \\ 0 \end{pmatrix}$

PROBLEM 8.4. Find the orthogonal complement to the linear vector space generated by the following vectors in $\mathbb{R}^4$

(a) $a_1 = \begin{pmatrix} 3 \\ 1 \\ -5 \\ 3 \end{pmatrix}$, $a_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 7 \end{pmatrix}$

(b) $a_1 = \begin{pmatrix} 1 \\ 1 \\ -6 \\ -7 \end{pmatrix}$, $a_2 = \begin{pmatrix} 1 \\ 2 \\ -12 \\ -8 \end{pmatrix}$

(c) $a_1 = \begin{pmatrix} 1 \\ -2 \\ -5 \\ 4 \end{pmatrix}$, $a_2 = \begin{pmatrix} 1 \\ 5 \\ 2 \\ 11 \end{pmatrix}$

(d) $a_1 = \begin{pmatrix} 4 \\ -3 \\ 18 \\ -5 \end{pmatrix}$, $a_2 = \begin{pmatrix} 2 \\ 1 \\ 4 \\ 5 \end{pmatrix}$

(e) $a_1 = \begin{pmatrix} 10 \\ 5 \\ 4 \\ 8 \end{pmatrix}$, $a_2 = \begin{pmatrix} 2 \\ 2 \\ -1 \\ 8 \end{pmatrix}$, $a_3 = \begin{pmatrix} 2 \\ -4 \\ 5 \\ 8 \end{pmatrix}$

(f) $a_1 = \begin{pmatrix} 7 \\ -8 \\ -4 \\ -13 \end{pmatrix}$, $a_2 = \begin{pmatrix} 7 \\ -8 \\ -5 \\ -16 \end{pmatrix}$, $a_3 = \begin{pmatrix} 3 \\ -1 \\ 5 \\ 17 \end{pmatrix}$

PROBLEM 8.5. Find the intersection of linear vector spaces $U_1 = \mathcal{L}(e_1, e_2, \ldots, e_k)$ and $U_2 = \mathcal{L}(f_1, f_2, \ldots, f_l)$ for the vector sets listed below.

(a) $a_1 = \begin{pmatrix} -9 \\ 0 \\ -9 \end{pmatrix}$, $a_2 = \begin{pmatrix} -7 \\ -8 \\ -3 \end{pmatrix}$,

$b_1 = \begin{pmatrix} -4 \\ -2 \\ -3 \end{pmatrix}$, $b_2 = \begin{pmatrix} -4 \\ 6 \\ 11 \end{pmatrix}$

(b) $a_1 = \begin{pmatrix} -1 \\ -3 \\ -3 \end{pmatrix}$, $a_2 = \begin{pmatrix} 5 \\ 5 \\ 3 \end{pmatrix}$,

$b_1 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$, $b_2 = \begin{pmatrix} -1 \\ -3 \\ -2 \end{pmatrix}$

(c) $a_1 = \begin{pmatrix} -3 \\ 1 \\ -3 \end{pmatrix}$, $a_2 = \begin{pmatrix} 1 \\ -3 \\ 1 \end{pmatrix}$,
$b_1 = \begin{pmatrix} -7 \\ 9 \\ -5 \end{pmatrix}$, $b_2 = \begin{pmatrix} 10 \\ 2 \\ 10 \end{pmatrix}$

(d) $a_1 = \begin{pmatrix} 1 \\ 9 \\ 7 \end{pmatrix}$, $a_2 = \begin{pmatrix} 2 \\ 6 \\ -1 \end{pmatrix}$,
$b_1 = \begin{pmatrix} 4 \\ -4 \\ 1 \end{pmatrix}$, $b_2 = \begin{pmatrix} 0 \\ -4 \\ -5 \end{pmatrix}$

(e) $a_1 = \begin{pmatrix} 1 \\ 4 \\ 5 \\ 4 \end{pmatrix}$, $a_2 = \begin{pmatrix} -1 \\ 0 \\ -3 \\ -4 \end{pmatrix}$, $a_3 = \begin{pmatrix} 4 \\ 5 \\ 9 \\ 6 \end{pmatrix}$,
$b_1 = \begin{pmatrix} 10 \\ 11 \\ 6 \\ -12 \end{pmatrix}$, $b_2 = \begin{pmatrix} 6 \\ 5 \\ 3 \\ -7 \end{pmatrix}$, $b_3 = \begin{pmatrix} 0 \\ -4 \\ -1 \\ 1 \end{pmatrix}$

(f) $a_1 = \begin{pmatrix} -2 \\ 4 \\ 0 \\ -4 \end{pmatrix}$, $a_2 = \begin{pmatrix} 1 \\ 6 \\ -1 \\ -3 \end{pmatrix}$, $a_3 = \begin{pmatrix} -4 \\ 4 \\ 2 \\ -3 \end{pmatrix}$,
$b_1 = \begin{pmatrix} 5 \\ -4 \\ -6 \\ -2 \end{pmatrix}$, $b_2 = \begin{pmatrix} 4 \\ 4 \\ 6 \\ 12 \end{pmatrix}$, $b_3 = \begin{pmatrix} -3 \\ 4 \\ -2 \\ -8 \end{pmatrix}$

(g) $a_1 = \begin{pmatrix} -8 \\ -4 \\ 1 \\ -1 \end{pmatrix}$, $a_2 = \begin{pmatrix} -8 \\ -7 \\ 1 \\ 5 \end{pmatrix}$, $a_3 = \begin{pmatrix} -4 \\ -8 \\ 4 \\ 8 \end{pmatrix}$,
$b_1 = \begin{pmatrix} 0 \\ 2 \\ 1 \\ 2 \end{pmatrix}$, $b_2 = \begin{pmatrix} -8 \\ -7 \\ 1 \\ 5 \end{pmatrix}$, $b_3 = \begin{pmatrix} -10 \\ -4 \\ -5 \\ -4 \end{pmatrix}$

(h) $a_1 = \begin{pmatrix} 2 \\ 4 \\ -2 \\ -6 \end{pmatrix}$, $a_2 = \begin{pmatrix} -7 \\ 4 \\ 4 \\ -4 \end{pmatrix}$, $a_3 = \begin{pmatrix} 4 \\ -1 \\ 5 \\ -1 \end{pmatrix}$,

$$b_1 = \begin{pmatrix} 10 \\ -1 \\ -4 \\ -6 \end{pmatrix}, \; b_2 = \begin{pmatrix} 11 \\ 1 \\ 0 \\ -10 \end{pmatrix}, \; b_3 = \begin{pmatrix} 4 \\ 2 \\ -3 \\ 1 \end{pmatrix}$$

# 9 Answers to problems

**[1.1]** (a) $x_1 = 2$, $x_2 = 2$, $x_3 = 1$, $x_4 = -5$. (b) $x_1 = 2$, $x_2 = -3$, $x_3 = -1$, $x_4 = -2$. (c) $x_1 = -2$, $x_2 = 0$, $x_3 = 1$, $x_4 = -4, x_5 = -4$. (d) $x_1 = 0$, $x_2 = -4$, $x_3 = 0$, $x_4 = 3$. (e) $x_1 = -1$, $x_2 = 2$, $x_3 = -2$, $x_4 = -2$. (f) $x_1 = 4$, $x_2 = -2$, $x_3 = 4$, $x_4 = 3$. (g) $x_1 = 0$, $x_2 = 4$, $x_3 = 1$, $x_4 = -2$. (h) $x_1 = 0$, $x_2 = 0$, $x_3 = -4$, $x_4 = -3$. (i) $x_1 = 3$, $x_2 = 0$, $x_3 = 1$, $x_4 = -4$. (j) $x_1 = -3$, $x_2 = -3$, $x_3 = 0$, $x_4 = -1$.

**[2.1]** No answer given.

**[2.2]** (a) $\lambda_1 = -16$, $\lambda_2 = -1$, $\lambda_3 = -14$. (b) $\lambda_1 = 3$, $\lambda_2 = -2$, $\lambda_3 = 0$. (c) $x_1 = 3$, $x_2 = -1$, $x_3 = -1$, $x_4 = 1$. (d) $x_1 = -1$, $x_2 = 1$, $x_3 = -2$, $x_4 = 2$. (e) $x_1 = 2$, $x_2 = -5$, $x_3 = -8$, $x_4 = -6$.

**[3.1]** (a) 1 (b) 1 (c) 4 (d) 2 (e) 1 (f) 2 (g) 2 (h) 1 (i) 2 (j) 2

**[3.2]**

$$\text{(a)}\ \ x = \begin{pmatrix} 0 \\ -3 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -2 \\ 3 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 6 \\ -7 \\ 0 \\ 1 \end{pmatrix}$$

$$\text{(b)}\ \ x = \begin{pmatrix} -5 \\ -5 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -4 \\ -3 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$

$$\text{(c)}\ \ x = \begin{pmatrix} -2 \\ 3 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2 \\ 2 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

$$\text{(d)}\ \ x = \begin{pmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 0 \\ 4 \\ -4 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -3 \\ -5 \\ 8 \\ 0 \\ 1 \end{pmatrix}$$

$$\text{(e)}\ \ x = \begin{pmatrix} 0 \\ -2 \\ 3 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 6 \\ -6 \\ -5 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 6 \\ -1 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$

(f) $x = \begin{pmatrix} -7 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 5 \\ -4 \\ 4 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 5 \\ -5 \\ 6 \\ 0 \\ 1 \end{pmatrix}$

(g) $x = \begin{pmatrix} 1 \\ 7 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -2 \\ -5 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 6 \\ 3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} -1 \\ 2 \\ 0 \\ 0 \\ 1 \end{pmatrix}$

(h) $x = \begin{pmatrix} 2 \\ -2 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2 \\ -4 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 5 \\ -7 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} -3 \\ 5 \\ 0 \\ 0 \\ 1 \end{pmatrix}$

(i) $x = \begin{pmatrix} 4 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -1 \\ 3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -7 \\ -4 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$

(j) $x = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -2 \\ -5 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} -7 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$

**[4.1]** (a) $-1152$ (b) 4 (c) $-96$ (d) $-288$ (e) $-48$ (f) 6 (g) 240 (h) 9 (i) $-4$ (j) 48 (k) 10 (l) 144

**[4.2]** (a) 2 (b) 2 (c) 4 (d) 2 (e) 2 (f) 3

**[4.3]** (a) $x_1 = 1$, $x_2 = -1$, $x_3 = 4$. (b) $x_1 = 3$, $x_2 = -2$, $x_3 = -3$. (c) $x_1 = -4$, $x_2 = 1$, $x_3 = -5$. (d) $x_1 = 2$, $x_2 = 4$, $x_3 = -2$, $x_4 = -5$. (e) $x_1 = -5$, $x_2 = 0$, $x_3 = 4$, $x_4 = 2$. (f) $x_1 = 1$, $x_2 = 3$, $x_3 = -2$. (g) $x_1 = 1$, $x_2 = -2$, $x_3 = 0$. (h) $x_1 = -1$, $x_2 = -2$, $x_3 = 1$, $x_4 = 1$.

**[5.1]** (a) $\begin{pmatrix} -2 & -3 & -11 & -7 & -12 \\ -9 & -9 & -12 & 5 & -5 \\ -1 & -1 & -5 & -3 & -5 \\ -1 & -1 & -1 & 1 & 0 \\ 2 & 3 & 2 & -2 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} -5 & 1 & 5 & -3 \\ -1 & 8 & 2 & -1 \\ 0 & 1 & 0 & 0 \\ -11 & -5 & 10 & -6 \end{pmatrix}$

(c) $\begin{pmatrix} 4 & 11 & -2 & 12 \\ 2 & 6 & -1 & 6 \\ -1 & 0 & 1 & 0 \\ -3 & -4 & 2 & -5 \end{pmatrix}$ (d) $\begin{pmatrix} -1 & 1 & 2 & -2 \\ -10 & -1 & 7 & -1 \\ 8 & 1 & -5 & 0 \\ 0 & -3 & -4 & 6 \end{pmatrix}$

(e) $\begin{pmatrix} -1 & 4 & -2 \\ -2 & 0 & -1 \\ 0 & 3 & -1 \end{pmatrix}$ (f) $\begin{pmatrix} 1 & -2 & 1 & -1 \\ 10 & -12 & 11 & -11 \\ 3 & -1 & 4 & -3 \\ 7 & -10 & 7 & -8 \end{pmatrix}$

(g) $\begin{pmatrix} 1 & -2 & 0 \\ 0 & -5 & 3 \\ -2 & 12 & -5 \end{pmatrix}$ (h) $\begin{pmatrix} -3 & -1 & 6 & 10 & -10 \\ -2 & -1 & 3 & 5 & -5 \\ -5 & -3 & 5 & 8 & -8 \\ 8 & 6 & -5 & -5 & 6 \\ -7 & -4 & 7 & 12 & -12 \end{pmatrix}$

(i) $\begin{pmatrix} -3 & -9 & -6 & 10 & -7 \\ -2 & 9 & 5 & -6 & -2 \\ 1 & 0 & 0 & -1 & 2 \\ -1 & -5 & -4 & 4 & -2 \\ 0 & 11 & 5 & -11 & 3 \end{pmatrix}$ (j) $\begin{pmatrix} -6 & -2 & 12 & -1 & -8 \\ 2 & 2 & 1 & 7 & 6 \\ 2 & 1 & -2 & 2 & 3 \\ -1 & -1 & 0 & -3 & -3 \\ -2 & 1 & 8 & 7 & 2 \end{pmatrix}$

**[5.2]** (a) $\begin{pmatrix} -1 & -12 & -10 \\ 12 & -12 & -28 \\ -8 & -1 & 10 \end{pmatrix}$ (b) $\begin{pmatrix} -3 & -6 & -4 & -6 \\ 1 & 23 & 36 & 57 \\ 27 & 45 & 18 & 26 \\ 2 & 17 & 23 & 37 \end{pmatrix}$

(c) $\begin{pmatrix} 4 & -17 & -26 & -22 & -20 \\ -1 & -14 & -14 & -18 & -14 \\ -8 & -39 & -27 & -37 & -25 \\ 9 & -48 & -70 & -61 & -55 \end{pmatrix}$ (d) $\begin{pmatrix} 2 & 4 & -3 \\ 2 & 16 & 1 \\ 2 & 9 & -3 \\ -1 & 7 & 6 \\ 4 & 3 & -8 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & -1 & -2 \\ 0 & 2 & 2 \\ 7 & -7 & -13 \end{pmatrix}$ (f) $\begin{pmatrix} -7 & 19 & 0 & -20 & -12 \\ -7 & 15 & -2 & -15 & -10 \\ 26 & -31 & 22 & 22 & 28 \\ 20 & -11 & 29 & -2 & 18 \end{pmatrix}$

**[6.1]** (a)$\lambda_1 = 7$, $\lambda_2 = 6$ (b)$\lambda_1 = 1$, $\lambda_2 = -1$ (c)$\lambda_1 = 0$, $\lambda_2 = -4$ (d)$\lambda_1 = -6$, $\lambda_2 = -5$ (e)$\lambda_1 = 1$, $\lambda_2 = -6$, $\lambda_3 = 2$ (f)$\lambda_1 = -2$, $\lambda_2 = -7$, $\lambda_3 = -6$ (g)$\lambda_1 = -7$, $\lambda_2 = 3$, $\lambda_3 = -5$ (h)$\lambda_1 = -4$, $\lambda_2 = -2$, $\lambda_3 = -2$, $\lambda_4 = -1$ (i)$\lambda_1 = -4$, $\lambda_2 = -1$, $\lambda_3 = -5$, $\lambda_4 = -2$ (j)$\lambda_1 = 6$, $\lambda_2 = -2$, $\lambda_3 = 2$

**[6.2]**

(a) $\begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix} \times \begin{pmatrix} 3 & -6 \\ 4 & -7 \end{pmatrix} \times \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -3 \end{pmatrix}$

(b) $\begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix} \times \begin{pmatrix} 5 & 4 \\ 0 & 7 \end{pmatrix} \times \begin{pmatrix} -1 & -2 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 7 \end{pmatrix}$

(c) $\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} \times \begin{pmatrix} -6 & -2 \\ 3 & -1 \end{pmatrix} \times \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} -4 & 0 \\ 0 & -3 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 6 & 14 \\ 0 & -8 \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & -8 \end{pmatrix}$

(e) $\begin{pmatrix} -1 & 0 & 1 \\ 1 & 1 & -1 \\ 0 & -1 & -1 \end{pmatrix} \times \begin{pmatrix} -2 & -6 & -2 \\ 2 & 2 & -2 \\ -2 & -6 & -2 \end{pmatrix} \times \begin{pmatrix} -2 & -1 & -1 \\ 1 & 1 & 0 \\ -1 & -1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -4 \end{pmatrix}$

(f) $\begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 3 & -8 & 3 \\ -3 & -2 & 3 \\ 2 & -8 & 4 \end{pmatrix} \times \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(g) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} \times \begin{pmatrix} -7 & 0 & -10 \\ 0 & 7 & 4 \\ 0 & 0 & 3 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} -7 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 7 \end{pmatrix}$

(h) $\begin{pmatrix} -2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix} \times \begin{pmatrix} -2 & 5 & -8 \\ 2 & 1 & -8 \\ -6 & 6 & -1 \end{pmatrix} \times \begin{pmatrix} 2 & -1 & 3 \\ 2 & -1 & 2 \\ 1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -4 \end{pmatrix}$

(i) $\begin{pmatrix} -71 & -65 & -6 \\ -1 & -2 & 2 \\ 12 & 11 & 1 \end{pmatrix} \times \begin{pmatrix} -4 & 2 & -2 \\ -1 & -7 & 2 \\ -1 & -2 & -3 \end{pmatrix} \times \begin{pmatrix} -24 & -1 & -142 \\ 25 & 1 & 148 \\ 13 & 1 & 77 \end{pmatrix} = \begin{pmatrix} -5 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -5 \end{pmatrix}$

(j) $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 3 & 1 & 3 \end{pmatrix} \times \begin{pmatrix} 3 & 7 & -3 \\ 3 & -1 & 3 \\ -5 & -7 & 1 \end{pmatrix} \times \begin{pmatrix} 3 & 1 & -1 \\ -3 & 0 & 1 \\ -2 & -1 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

(k) $\begin{pmatrix} -1 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 3 & -2 & 3 & -2 \\ -2 & 6 & 0 & 2 \\ 2 & -8 & -2 & -2 \\ 3 & -6 & 3 & -2 \end{pmatrix} \times \begin{pmatrix} 1 & 1 & -1 & -1 \\ -1 & 0 & 0 & 1 \\ 1 & 0 & 1 & -1 \\ 3 & 1 & -1 & -2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$

(l) $\begin{pmatrix} -2 & -3 & 3 & 2 \\ 5 & -1 & -6 & -13 \\ 1 & 0 & -1 & -2 \\ -1 & -1 & 1 & 1 \end{pmatrix} \times \begin{pmatrix} 0 & 6 & 1 & 8 \\ -8 & -8 & 1 & 0 \\ -6 & 0 & -7 & -2 \\ 3 & 3 & 4 & 5 \end{pmatrix} \times \begin{pmatrix} 0 & -1 & 7 & 1 \\ 1 & 1 & -8 & -5 \\ 2 & 1 & -8 & -7 \\ -1 & -1 & 7 & 4 \end{pmatrix} =$

$= \begin{pmatrix} -6 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 \end{pmatrix}$

**[7.1]**

(a) $2x_1^2 + x_2^2 - 5x_3^2$

(b) $-5x_1^2 - 2x_2^2 + x_3^2$

(c) $-5x_1^2 - 3x_2^2 + 3x_3^2$

(d) $-x_1^2 - 2x_2^2 + 2x_3^2$

(e) $2x_1^2 - 3x_2^2 + 3x_3^2 - 5x_4^2$

(f) $-5x_1^2 - 4x_2^2 + x_3^2 + 3x_4^2$

(g) $-3x_1^2 + 2x_2^2 + 3x_3^2 + 2x_4^2$

(h) $x_1^2 - 3x_2^2 + 4x_3^2 + x_4^2$

**[7.2]** (a) $n_+ = 2$, $n_- = 0$, positive semidefinite (b) $n_+ = 1$, $n_- = 2$, not definite (c) $n_+ = 2$, $n_- = 1$, not definite (d) $n_+ = 2$, $n_- = 1$, not definite (e) $n_+ = 1$, $n_- = 3$, not definite (f) $n_+ = 1$, $n_- = 3$, not definite (g) $n_+ = 0$, $n_- = 4$, negative definite (h) $n_+ = 0$, $n_- = 3$, negative semidefinite

**[8.1]**

(a) $f_1 = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$, $f_2 = \begin{pmatrix} -23 \\ -1 \\ 8 \end{pmatrix}$, $f_3 = \begin{pmatrix} -1 \\ 7 \\ -2 \end{pmatrix}$

(b) $f_1 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$, $f_2 = \begin{pmatrix} -2 \\ 1 \\ 4 \end{pmatrix}$, $f_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$

(c) $f_1 = \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix}$, $f_2 = \begin{pmatrix} 4 \\ 31 \\ -37 \end{pmatrix}$, $f_3 = \begin{pmatrix} -8 \\ 7 \\ 5 \end{pmatrix}$

(d) $f_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $f_2 = \begin{pmatrix} -4 \\ 8 \\ 5 \end{pmatrix}$, $f_3 = \begin{pmatrix} -1 \\ 2 \\ -4 \end{pmatrix}$

(e) $f_1 = \begin{pmatrix} 2 \\ 2 \\ 3 \\ 1 \end{pmatrix}$, $f_2 = \begin{pmatrix} -2 \\ 4 \\ -3 \\ 5 \end{pmatrix}$, $f_3 = \begin{pmatrix} 43 \\ 4 \\ -30 \\ -4 \end{pmatrix}$

(f) $f_1 = \begin{pmatrix} 3 \\ 4 \\ 1 \\ 2 \end{pmatrix}$, $f_2 = \begin{pmatrix} 9 \\ -14 \\ 13 \\ 8 \end{pmatrix}$, $f_3 = \begin{pmatrix} 9 \\ 3 \\ 13 \\ -26 \end{pmatrix}$

(g) $f_1 = \begin{pmatrix} 3 \\ 1 \\ 3 \\ 1 \\ 1 \end{pmatrix}$, $f_2 = \begin{pmatrix} 15 \\ -2 \\ -13 \\ 12 \\ -16 \end{pmatrix}$, $f_3 = \begin{pmatrix} 12 \\ -13 \\ -18 \\ 2 \\ 29 \end{pmatrix}$, $f_4 = \begin{pmatrix} 7 \\ -26 \\ 9 \\ -14 \\ -8 \end{pmatrix}$

(h) $f_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 3 \\ 3 \end{pmatrix}$, $f_2 = \begin{pmatrix} 3 \\ -4 \\ -3 \\ -1 \\ 1 \end{pmatrix}$, $f_3 = \begin{pmatrix} 3 \\ 10 \\ -12 \\ 4 \\ -1 \end{pmatrix}$, $f_4 = \begin{pmatrix} -30 \\ -16 \\ -15 \\ 17 \\ -2 \end{pmatrix}$

**[8.2]**

(a) $\begin{pmatrix} 11 & 2 & 8 \\ 2 & 6 & -5 \\ 8 & -5 & 14 \end{pmatrix}$

(b) $\begin{pmatrix} 12 & 0 & -4 \\ 0 & 2 & 1 \\ -4 & 1 & 6 \end{pmatrix}$

(c) $\begin{pmatrix} 11 & 4 & 3 \\ 4 & 14 & 2 \\ 3 & 2 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} 6 & 5 & 2 \\ 5 & 17 & 5 \\ 2 & 5 & 11 \end{pmatrix}$

(e) $\begin{pmatrix} 21 & -4 & -9 & -11 \\ -4 & 13 & 3 & 13 \\ -9 & 3 & 19 & 7 \\ -11 & 13 & 7 & 28 \end{pmatrix}$

(f) $\begin{pmatrix} 30 & -6 & -1 & 2 \\ -6 & 15 & -7 & 8 \\ -1 & -7 & 5 & -1 \\ 2 & 8 & -1 & 23 \end{pmatrix}$

(g) $\begin{pmatrix} 14 & 2 & 3 & -19 \\ 2 & 28 & 16 & 2 \\ 3 & 16 & 10 & -3 \\ -19 & 2 & -3 & 34 \end{pmatrix}$

(h) $\begin{pmatrix} 22 & -3 & -8 & 6 \\ -3 & 29 & 4 & -15 \\ -8 & 4 & 9 & 0 \\ 6 & -15 & 0 & 11 \end{pmatrix}$

**[8.3]**

(a) $p = \begin{pmatrix} -1 \\ -2 \\ 2 \end{pmatrix}$, $d^2 = 32$

(b) $p = \begin{pmatrix} -5 \\ -3 \\ 3 \end{pmatrix}$, $d^2 = 26$

(c) $p = \begin{pmatrix} 2 \\ 2 \\ -4 \end{pmatrix}$, $d^2 = 3$

(d) $p = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}$, $d^2 = 27$

(e) $p = \begin{pmatrix} 3 \\ 4 \\ 1 \\ -1 \end{pmatrix}$, $d^2 = 18$

(f) $p = \begin{pmatrix} 1 \\ -2 \\ -5 \\ -1 \end{pmatrix}$, $d^2 = 24$

(g) $p = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 1 \end{pmatrix}$, $d^2 = 18$

(h) $p = \begin{pmatrix} 5 \\ -4 \\ -3 \\ 0 \end{pmatrix}$, $d^2 = 16$

**[8.4]**

(a) $b_1 = \begin{pmatrix} 3 \\ -4 \\ 1 \\ 0 \end{pmatrix}$, $b_2 = \begin{pmatrix} 2 \\ -9 \\ 0 \\ 1 \end{pmatrix}$

(b) $b_1 = \begin{pmatrix} 0 \\ 6 \\ 1 \\ 0 \end{pmatrix}$, $b_2 = \begin{pmatrix} 6 \\ 1 \\ 0 \\ 1 \end{pmatrix}$

(c) $b_1 = \begin{pmatrix} 3 \\ -1 \\ 1 \\ 0 \end{pmatrix}$, $b_2 = \begin{pmatrix} -6 \\ -1 \\ 0 \\ 1 \end{pmatrix}$

(d) $b_1 = \begin{pmatrix} -3 \\ 2 \\ 1 \\ 0 \end{pmatrix}$, $b_2 = \begin{pmatrix} -1 \\ -3 \\ 0 \\ 1 \end{pmatrix}$

(e) $b = \begin{pmatrix} -8 \\ 8 \\ 8 \\ 1 \end{pmatrix}$

(f) $b = \begin{pmatrix} -1 \\ -1 \\ -3 \\ 1 \end{pmatrix}$

**[8.5]**

(a) $c = \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix}$ (b) $c = \begin{pmatrix} -1 \\ -3 \\ -3 \end{pmatrix}$ (c) $c = \begin{pmatrix} -5 \\ -1 \\ -5 \end{pmatrix}$ (d) $c = \begin{pmatrix} 0 \\ -4 \\ -5 \end{pmatrix}$

(e) $c_1 = \begin{pmatrix} -2 \\ -1 \\ -1 \\ 2 \end{pmatrix}$, $c_2 = \begin{pmatrix} -2 \\ -3 \\ -2 \\ 2 \end{pmatrix}$

(f) $c_1 = \begin{pmatrix} 4 \\ -4 \\ -2 \\ 3 \end{pmatrix}$, $c_2 = \begin{pmatrix} -3 \\ -2 \\ 1 \\ -1 \end{pmatrix}$

(g) $c_1 = \begin{pmatrix} 2 \\ 4 \\ -2 \\ -4 \end{pmatrix}$, $c_2 = \begin{pmatrix} 4 \\ -1 \\ 3 \\ 3 \end{pmatrix}$

(h) $c_1 = \begin{pmatrix} 4 \\ -1 \\ -5 \\ 1 \end{pmatrix}$, $c_2 = \begin{pmatrix} -5 \\ -1 \\ 1 \\ 3 \end{pmatrix}$

# 10 Sample Midterm Exam I

## SECTION I: MULTIPLE CHOICE

**Directions:** Each of the following problems is followed by five choices. Select the best choice and put the corresponding mark in your answer sheet. **Calculators may NOT be used at any part of the exam.**

1. Which of the following statements are true for any pair of non-degenerate $n \times n$ matrices $A$ and $B$?

   I. $(AB)^T = A^T B^T$
   II. $(AB)^T = B^T A^T$
   III. $(A^T)^{-1} = (A^{-1})^T$

   **(A)** I only
   **(B)** II only
   **(C)** I and III
   **(D)** II and III
   **(E)** None of the above options

2. Let $A = (a_{ij})$ be a 5×5 matrix such that $a_{11} = 1$, $a_{22} = 2$, and $\det A = 0$. The largest (by absolute value) third-order minor of $A$ is

   **(A)** 0
   **(B)** 1
   **(C)** 2
   **(D)** 3
   **(E)** Could be any number

3. Let $\{a, b\}$ and $\{c, d\}$ be two sets of vectors, both linearly independent. Which of the following MUST be true?

   I. $\{a, b, c, d\}$ is also linearly independent.
   II. Both $\{a, c\}$ and $\{b, d\}$ are linearly independent
   III. Either $\{a, c\}$ or $\{b, d\}$ is linearly independent

   **(A)** I
   **(B)** II
   **(C)** III
   **(D)** All are true
   **(E)** Neither is true

4. Let $A$ be a matrix such that $A \cdot B = 0$ for any compartible matrix $B$. Which of the following MUST be true?

   I. $\det A = 0$

   II. $A = 0$

   III. Rows of $A$ are linearly dependent

   **(A)** I
   **(B)** II
   **(C)** III
   **(D)** I & III
   **(E)** II & III

5. Let $(x_1, x_2, x_3)$ be a solution to the following SLE

$$\begin{cases} 7x_1 + 5x_2 = 0 \\ 3x_1 + x_2 + x_3 = -5 \\ 3x_1 + 2x_2 - x_3 = 5 \end{cases}$$

   Then $x_3$ is
   **(A)** 3
   **(B)** -3
   **(C)** 5
   **(D)** -5
   **(E)** None of these

6. The determinant $\begin{vmatrix} -1 & -12 & -8 \\ 0 & 4 & 8 \\ 0 & 8 & -8 \end{vmatrix}$ is equal to
   **(A)** 0
   **(B)** 4
   **(C)** 52
   **(D)** 64
   **(E)** None of these

7. Let $A$ be the matrix that is inverse to $\begin{pmatrix} 6 & -5 & -2 \\ -1 & 0 & -1 \\ -2 & 2 & 1 \end{pmatrix}$. Then $a_{22} =$
   **(A)** 1
   **(B)** 2
   **(C)** 3
   **(D)** 5
   **(E)** None of these

8. Let $A$ be a $3\times4$ matrix and $B$ be a $4\times5$ matrix. $\mathrm{rk}(A\cdot B)$ is not greater than
   **(A)** 0
   **(B)** 3
   **(C)** 4
   **(D)** 5
   **(E)** Could be any number

9. Let $A$, $B$, and $C$ be square matrices such that $A\cdot B = E$ and $B\cdot C = E$, where $E$ is the identity matrix. Which of the following MUST be true?

   I. $\det(A) = \det(C)$

   II. $A = C$

   III. If $A = B$ then $|\det(A)| = 1$

   **(A)** I
   **(B)** II
   **(C)** III
   **(D)** I & II
   **(E)** I, II, and & III

10. The rank of the matrix $\begin{pmatrix} 2 & 0 & 12 & 6 \\ -3 & 0 & 2 & 3 \\ -7 & -4 & -4 & -1 \end{pmatrix}$ is
   **(A)** 0
   **(B)** 1
   **(C)** 2
   **(D)** 3
   **(E)** 4

## SECTION II: FREE RESPONSE

**Directions:** Show all your work. Indicate clearly the methods you use because you will be graded on the correctness of your methods as well as on the accuracy of your results and explanations. *You may use the back of this sheet for writing.*

1. Evaluate the following determinant using row or column decomposition.

$$\begin{vmatrix} 12 & 4 & 12 & -2 \\ -7 & 10 & 4 & 1 \\ 6 & -3 & 16 & -1 \\ 13 & -18 & 16 & -2 \end{vmatrix}$$

2. Solve the following matrix equation. Explain all steps in your calculations.

$$\begin{pmatrix} -9 & -11 & 10 \\ 9 & 10 & -9 \\ 8 & 9 & -8 \end{pmatrix} X = \begin{pmatrix} -3 & 4 \\ -5 & -4 \\ 1 & 1 \end{pmatrix}$$

3. Find the fundamental set of solutions to the following system of linear equations. Indicate the dimension of the solution space.

$$\begin{cases} x_1 + 2x_2 - 2x_3 - 17x_4 - 3x_5 = 0 \\ 3x_1 - 5x_2 + 2x_3 - 11x_4 - x_5 = 14 \\ x_1 - 4x_2 - 7x_4 - x_5 = 10 \end{cases}$$

# ANSWERS

**Multiple Choice:**

1. D
2. E
3. E
4. E
5. D
6. E
7. B
8. B
9. E
10. D

**Free Response:**

1. det $= -60$
2. $\begin{pmatrix} -14 & -5 \\ 49 & 41 \\ 41 & 41 \end{pmatrix}$
3. For instance, $x = \begin{pmatrix} 2 \\ -2 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 7 \\ 0 \\ -5 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}$
   dim $= 2$
   Note: the answer is not unique.

# 11 Sample Midterm Exam II

## SECTION I: MULTIPLE CHOICE

**Directions:** Each of the following problems is followed by five choices. Select the best choice and put the corresponding mark in your answer sheet. **Calculators may NOT be used at any part of the exam.**

1. Which of the following sets form linear vector spaces with respect to the function addition operation?

   I. $\{y = f(x) \mid f'(0) = 1\}$
   II. $\{y = f(x) \mid f'(1) = 0\}$
   III. $\{y = f(x) \mid f'(0) = f'(1)\}$

   **(A)** I
   **(B)** II
   **(C)** I & II
   **(D)** II & III
   **(E)** I, II, and III

2. Let $A = (a_{ij})$ be a 4×4 matrix such that $a_{11} = 1$, $a_{22} = 2$, and $\det A = 0$. The smallest (by absolute value) third-order minor of $A$ is

   **(A)** 0
   **(B)** 1
   **(C)** 2
   **(D)** 3
   **(E)** Could be any number

3. Which of the following sets form linear vector spaces with respect to matrix addition?

   I. The set of all 2×2 matrices that are equal to their transpose
   II. The set of all 3×3 matrices with the property $\det A = 0$
   III. The set of all 4×4 matrices with the property $\det A = 1$.

   **(A)** I
   **(B)** II
   **(C)** III
   **(D)** I & II
   **(E)** I and III

4. Let $A = (a_{ij})$ be a $3 \times 4$ matrix such that $a_{11} = 0$. Whe rank of $A$ could be each of the following EXCEPT

**(A)** 0
**(B)** 1
**(C)** 2
**(D)** 3
**(E)** 4

5. Let $a$, $b$, and $c$ be vectors in $\mathbb{R}^3$ that are linearly independent. Which of the following MUST be true?

I. Vectors $a$ and $b$ also form a linearly independent set of vectors
II. The set of vectors $\{a, b, c\}$ is full in $\mathbb{R}^3$
III. The set $\{a, b, c, d\}$ is linearly dependent for any vector $d \in \mathbb{R}^3$

**(A)** I
**(B)** II
**(C)** I & II
**(D)** I & III
**(E)** I, II and III

6. Let $A$ and $B$ be non-degenerate $6 \times 6$ matrices. Which of the following MUST be true?

I. $(A \cdot B)^{-1} = A^{-1} \cdot B^{-1}$
II. $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$
III. The transpose to $A^{-1}$ is the inverse to $A^{\mathrm{T}}$.

**(A)** I
**(B)** II
**(C)** I & III
**(D)** I & II
**(E)** II & III

7. Let $(x_1, x_2, x_3)$ be a solution to the following SLE

$$\begin{cases} 2x_1 - 7x_2 - 5x_3 = 4 \\ x_1 + x_2 + 2x_3 = -7 \\ 6x_1 - 5x_2 + 5x_3 = -12 \end{cases}$$

Then $x_1$ is
**(A)** -7

**(B)** -4
**(C)** 0
**(D)** 7
**(E)** None of these

8. The determinant $\begin{vmatrix} -3 & 0 & 12 \\ 4 & 1 & 1 \\ -11 & -2 & 1 \end{vmatrix}$ is equal to

**(A)** 0
**(B)** 10
**(C)** 18
**(D)** 27
**(E)** None of these

9. If $A$ is the matrix inverse to $\begin{pmatrix} -5 & -2 & 2 \\ 9 & 5 & -3 \\ 4 & 1 & -2 \end{pmatrix}$ then $a_{13} =$

**(A)** 4
**(B)** -4
**(C)** 11
**(D)** -11
**(E)** None of these

10. The rank of the matrix $\begin{pmatrix} -10 & 0 & -6 & -4 \\ 5 & -10 & -1 & -4 \\ -5 & 10 & 1 & 4 \end{pmatrix}$ is

**(A)** 0
**(B)** 1
**(C)** 2
**(D)** 3
**(E)** 4

## SECTION II: FREE RESPONSE

**Directions:** Show all your work. Indicate clearly the methods you use because you will be graded on the correctness of your methods as well as on the accuracy of your results and explanations. *You may use the back of this sheet for writing.*

1. Evaluate the following determinant using row or column decomposition.

$$\begin{vmatrix} 10 & 5 & -18 & -4 \\ 11 & 5 & -17 & -2 \\ -2 & 0 & 13 & 3 \\ 11 & 10 & 10 & 1 \end{vmatrix}$$

2. Solve the following matrix equation. Explain all steps in your calculations.

$$X \begin{pmatrix} 3 & 11 & -11 \\ -1 & -2 & 3 \\ -1 & -4 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 4 & -3 \\ -5 & -3 & -2 \end{pmatrix}$$

3. Find the fundamental set of solutions to the following system of linear equations. Indicate the dimension of the solution space.

$$\begin{cases} x_1 + 5x_2 + 5x_3 + x_4 - 5x_5 = -1 \\ x_1 - 5x_2 - 7x_3 + 13x_4 + x_5 = 3 \\ 2x_2 + 5x_3 - 5x_4 - 9x_5 = -6 \end{cases}$$

# ANSWERS

**Multiple Choice:**

1. D
2. A
3. A
4. E
5. E
6. E
7. A
8. D
9. B
10. C

**Free Response:**

1. det $= -100$
2. $\begin{pmatrix} 6 & 1 & 15 \\ -27 & -5 & -71 \end{pmatrix}$
3. For instance, $x = \begin{pmatrix} -1 \\ 2 \\ -2 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -6 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 5 \\ -3 \\ 3 \\ 0 \\ 1 \end{pmatrix}$
   dim $= 2$
   Note: the answer is not unique.

# 12 Sample Midterm Exam III

## SECTION I: MULTIPLE CHOICE

**Directions:** Each of the following problems is followed by five choices. Select the best choice and put the corresponding mark in your answer sheet. **Calculators may NOT be used at any part of the exam.**

1. Let $A$ be the inverse matrix for $\begin{pmatrix} -4 & 3 & 5 \\ -1 & 1 & 2 \\ -3 & 3 & 5 \end{pmatrix}$. Then $a_{13} =$

   **(A)** 1
   **(B)** 0
   **(C)** -1
   **(D)** 3
   **(E)** None of these

2. Let $A$ and $B$ be two $7\times 7$ matrices such that $\mathrm{rk}(A) = 2$ and $\mathrm{rk}(B) = 3$. Which of the following COULD be true?

   I. $\mathrm{rk}(A + B) = 0$
   II. $\mathrm{rk}(A + B) = 1$
   III. $\mathrm{rk}(A + B) = 6$

   **(A)** I only
   **(B)** II only
   **(C)** I and II
   **(D)** I, II, and III
   **(E)** None

3. The determinant $\begin{vmatrix} 7 & 1 & 4 \\ 6 & 0 & -6 \\ 10 & 3 & 15 \end{vmatrix}$ is equal to

   **(A)** 6
   **(B)** 12
   **(C)** 36
   **(D)** 48
   **(E)** None of these

4. Let $a$, $b$, $c$, and $d$ be four vectors such that $\{a, b, c\}$ is linearly dependent and $\{b, c, d\}$ is also linearly dependent. Which of the following MUST be true?

   I. $\{a, b, d\}$ is linearly dependent
   II. $\dim \mathcal{L}(a, b, c, d) \leq 2$

III. $d - a$ can be expressed as a linear combination of $b$ and $c$

**(A)** I only
**(B)** I and II
**(C)** I, II, and III
**(D)** III only
**(E)** None

5. The rank of the following matrix is equal to

$$\begin{pmatrix} -3 & 2 & 2 & 1 & -4 \\ 9 & -6 & -6 & -3 & 12 \\ -15 & 10 & 10 & 5 & -20 \end{pmatrix}$$

**(A)** 1
**(B)** 2
**(C)** 3
**(D)** 4
**(E)** 5

6. Let $a$, $b$, $c$, and $d$ be four vectors such that $\{b, c\}$ is linearly independent, $\{a, b, c\}$ is linearly dependent, and $\{b, c, d\}$ is also linearly dependent. Which of the following MUST be true?

I. $\{a, b, d\}$ is linearly dependent

II. $\dim \mathcal{L}(a, b, c, d) \leq 2$

III. $d - a$ can be expressed as a linear combination of $b$ and $c$

**(A)** I
**(B)** I and II
**(C)** I, II, and III
**(D)** III only
**(E)** None

7. If $A = \begin{pmatrix} -1 & 4 \\ 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix}$, and $C = B^T \times A$, then $c_{21} =$

**(A)** 1
**(B)** 9
**(C)** -3
**(D)** 12
**(E)** -1

8. The dimension of the space of solutions of $x_1 + 3x_2 + 15x_3 + 12x_4 + 15x_5 = 0$ is

**(A)** 1
**(B)** 2
**(C)** 3

**(D)** 4
**(E)** 5

9. Let $(x_1, x_2, x_3)$ be a solution to the following SLE

$$\begin{cases} 8x_1 + 7x_2 - 3x_3 = -2 \\ 9x_1 + 3x_2 + 4x_3 = -12 \\ x_1 - 3x_2 - 5x_3 = -8 \end{cases}$$

Then $x_1 + x_2 + x_3 =$
**(A)** -2
**(B)** 0
**(C)** 2
**(D)** 10
**(E)** None of these

10. Which of the following statements are true for any pair of non-degenerate $n \times n$ matrices $A$ and $B$?

I. $(B^{-1}AB)^{-1} = B^{-1}A^{-1}B$
II. $(A^TB^T)^{-1} = (A^{-1}B^{-1})^T$
III. $(A^TB^T)^{-1} = (B^{-1}A^{-1})^T$

**(A)** I
**(B)** I and II
**(C)** I and III
**(D)** III only
**(E)** None is true

## SECTION II: FREE RESPONSE

**Directions:** Show all your work. Indicate clearly the methods you use because you will be graded on the correctness of your methods as well as on the accuracy of your results and explanations. *You may use the back of this sheet for writing.*

1.
$$\begin{vmatrix} -3 & 0 & -7 & -1 & 7 \\ -6 & 0 & 0 & 0 & 0 \\ -7 & 0 & -1 & 0 & 4 \\ 9 & -1 & 0 & 5 & 4 \\ 2 & -1 & 5 & 6 & -2 \end{vmatrix}$$

2. Solve the following matrix equation. Explain each step in your solution.
$$\begin{pmatrix} -3 & 1 & -1 \\ 0 & 1 & 1 \\ -1 & -2 & -3 \end{pmatrix} X = \begin{pmatrix} -2 & 0 \\ 2 & -6 \\ -7 & 1 \end{pmatrix}$$

3. Find the fundamental set of solutions to the following system of linear equations. Indicate the dimension of the solution space.
$$\begin{cases} 5x_1 - 7x_2 - 8x_3 - 7x_4 - 21x_5 + 7x_6 = -10 \\ 3x_1 - 4x_2 - 4x_3 + 2x_4 - 10x_5 + 6x_6 = -3 \\ 4x_1 - 5x_2 - 3x_3 + 22x_4 - 5x_5 + 14x_6 = 7 \end{cases}$$

# ANSWERS

**Multiple Choice:**

1. A
2. B
3. D
4. E
5. A
6. C
7. C
8. D
9. B
10. B

**Free Response:**

1. $\det = -42$
2. $X = \begin{pmatrix} -2 & -28 \\ -3 & -45 \\ 5 & 39 \end{pmatrix}$
3. $x = \begin{pmatrix} -5 \\ -9 \\ 6 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -6 \\ 5 \\ -9 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 3 \\ -4 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} -2 \\ 3 \\ -3 \\ 0 \\ 0 \\ 1 \end{pmatrix}$

# 13 Sample Final Exam I

## SECTION I: MULTIPLE CHOICE

**Directions:** Each of the following problems is followed by five choices. Select the best choice and put the corresponding mark in your answer sheet. **Calculators may NOT be used at any part of the exam.**

1. Let $A$ be a 3×3 matrix and let $B = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. If $A$ is multiplied by $B$ from the left then

   **(A)** The 1$^{\text{st}}$ column of $A$ is multipled by 5
   **(B)** The 1$^{\text{st}}$ row of $A$ is multipled by 5
   **(C)** The 1$^{\text{st}}$ column of $A$ is divided by 5
   **(D)** The 1$^{\text{st}}$ row of $A$ is divided by 5
   **(E)** None of these

2. The signature of the quadratic form $-10x_1^2 - 6x_1x_2 + 6x_1x_3 - x_2^2 + 2x_2x_3 - x_3^2$ is

   **(A)** -2
   **(B)** -1
   **(C)** 0
   **(D)** 1
   **(E)** 2

3. Let $A$ be a 3×3 matrix with eigenvalues $\lambda_1$, $\lambda_2$, and $\lambda_3$ such that $\lambda_1 \neq \lambda_2$, $\lambda_1 \neq \lambda_3$, and $\lambda_2 \neq \lambda_3$. Which of the following MUST be true?

   I If $a$ and $b$ are eigenvectors corresponding to $\lambda_1$ then the set $\{a, b\}$ is linearly dependent

   II. If $a_1$ and $a_2$ are eigenvectors corresponding to $\lambda_1$ and $\lambda_2$, respectively, then $\{a_1, a_2\}$ is linearly independent

   III. $\det(A) \neq 0$

   **(A)** I
   **(B)** II
   **(C)** III
   **(D)** I and II
   **(E)** II and III

4. Which of the following could be the set of eigenvalues of $\begin{pmatrix} -2 & -4 & -4 \\ 2 & -2 & 2 \\ -2 & 4 & 0 \end{pmatrix}$?

**(A)** $\lambda_1 = 2$, $\lambda_2 = -4$, $\lambda_3 = -2$
**(B)** $\lambda_1 = 2$, $\lambda_2 = -4$, $\lambda_3 = 2$
**(C)** $\lambda_1 = 2$, $\lambda_2 = 4$, $\lambda_3 = -2$
**(D)** $\lambda_1 = 2$, $\lambda_2 = 4$, $\lambda_3 = 2$
**(E)** None of these

5. The quadratic form $8x_1^2 + 16x_1x_2 + 16x_1x_3 + 4x_2^2 + 8x_2x_3 + 8x_3^2$ is

**(A)** positively definite
**(B)** negatively definite
**(C)** non-positively definite
**(D)** non-negatively definite
**(E)** not definite

6. Let $d$ be the distance between the endpoint of vector $x$ and the linear space generated by vectors $x = \begin{pmatrix} 13 \\ -1 \\ 3 \end{pmatrix}$, $e_1 = \begin{pmatrix} -2 \\ -2 \\ -4 \end{pmatrix}$, $e_2 = \begin{pmatrix} 4 \\ -3 \\ 1 \end{pmatrix}$ Then, $d^2 =$

**(A)** 1
**(B)** 16
**(C)** 27
**(D)** 32
**(E)** None of these

7. Let $A$ be a 3×3 matrix. Which of the following MIGHT be true?

I. $A$ has no real eigenvalues
II. $A$ is diagonalizable but $A^2$ is not
III. $A^2$ is diagonalizable but $A$ is not

**(A)** I
**(B)** II
**(C)** III
**(D)** I and II
**(E)** I and III

8. Which of the following is NOT an eigenvalue of $A = \begin{pmatrix} -3 & -3 & 1 & -1 \\ 0 & -5 & 1 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & -3 & 3 & -2 \end{pmatrix}$?

**(A)** -2

**(B)** -3
**(C)** -4
**(D)** -5
**(E)** -6

9. The element $g_{12}$ of the Gram matrix of $e_1 = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$, $e_2 = \begin{pmatrix} -4 \\ 2 \\ 1 \end{pmatrix}$, $e_3 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ is

**(A)** 5
**(B)** 6
**(C)** -5
**(D)** 21
**(E)** None of these

10. Let $A$ be a $3\times 3$ matrix with the property $A = A^2$. Which of the following MUST be true?

I $\det(A) = 0$

II. $\det(A) = 1$

III. $A$ is the identity matrix

**(A)** I
**(B)** II
**(C)** III
**(D)** II and III
**(E)** None is true

## SECTION II: FREE RESPONSE

**Directions:** Show all your work. Indicate clearly the methods you use because you will be graded on the correctness of your methods as well as on the accuracy of your results and explanations. *You may use the back of this sheet for writing.*

1. [**6 pts**] For the following matrix $A$ find matrices $C$ and $C^{-1}$ such that $C^{-1}AC$ is diagonal and write a product expression for $e^A$. Show your work.

$$A = \begin{pmatrix} -1 & 3 & -1 \\ 2 & -5 & -2 \\ 2 & 3 & -4 \end{pmatrix}$$

2. [**4 pts**] Use Gram-Schmidt orthogonalization process to find an orthogonal base in the linear hull of the following system of vectors. Show your work.

$$e_1 = \begin{pmatrix} 1 \\ 4 \\ 3 \\ 3 \end{pmatrix}, e_2 = \begin{pmatrix} -3 \\ 4 \\ -1 \\ -1 \end{pmatrix}, e_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

3. [**4 pts**] For the vectors $x$, $e_1$, $e_2$, and $e_3$ below, find the projection of the vector $x$ onto the linear hull of $e_1$, $e_2$, and $e_3$. Evaluate the distance between the endpoint of $x$ and the linear hull.

$$x = \begin{pmatrix} -4 \\ -14 \\ 0 \\ -6 \end{pmatrix}, e_1 = \begin{pmatrix} 3 \\ -4 \\ 3 \\ -5 \end{pmatrix}, e_2 = \begin{pmatrix} -2 \\ 0 \\ -1 \\ 5 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ -2 \\ -1 \\ 3 \end{pmatrix}$$

# ANSWERS

**Multiple Choice:**

1. B
2. B
3. D
4. A
5. E
6. C
7. C
8. E
9. A
10. E

**Free Response:**

1. $\begin{pmatrix} -4 & -1 & 3 \\ -1 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix} \times \begin{pmatrix} -1 & 3 & -1 \\ 2 & -5 & -2 \\ 2 & 3 & -4 \end{pmatrix} \times \begin{pmatrix} -1 & 4 & -1 \\ 0 & -1 & 1 \\ -1 & 5 & -1 \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -5 \end{pmatrix}$

2. $f_1 = \begin{pmatrix} 1 \\ 4 \\ 3 \\ 3 \end{pmatrix}$, $f_2 = \begin{pmatrix} -2 \\ 2 \\ -1 \\ -1 \end{pmatrix}$, $f_3 = \begin{pmatrix} 8 \\ 4 \\ 3 \\ -11 \end{pmatrix}$

3. $p = \begin{pmatrix} 4 \\ -10 \\ 4 \\ -2 \end{pmatrix}$, $d^2 = 112$

# 14 Sample Final Exam II

## SECTION I: MULTIPLE CHOICE

**Directions:** Each of the following problems is followed by five choices. Select the best choice and put the corresponding mark in your answer sheet. **Calculators may NOT be used at any part of the exam.**

1. Let $A$ be a $3\times3$ matrix and let $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. If $A$ is multiplied by $B$ from the left then

   **(A)** The $2^{\text{nd}}$ column of $A$ is multipled by 3
   **(B)** The $2^{\text{nd}}$ row of $A$ is multipled by 3
   **(C)** The $2^{\text{nd}}$ column of $A$ is divided by 3
   **(D)** The $2^{\text{nd}}$ row of $A$ is divided by 3
   **(E)** None of these

2. The signature of the quadratic form $4x_1^2 - 8x_1x_2 + 16x_1x_3 + 9x_2^2 - 16x_2x_3 + 10x_3^2$ is

   **(A)** -2
   **(B)** -1
   **(C)** 0
   **(D)** 1
   **(E)** 2

3. Let $A$ be a $3\times3$ matrix with eigenvalues $\lambda_1$, $\lambda_2$, and $\lambda_3$ such that $\lambda_1 \neq \lambda_2$, $\lambda_1 \neq \lambda_3$, and $\lambda_2 \neq \lambda_3$. Which of the following MUST be true?

   I There exists a basis of $\mathbb{R}^3$ that consists of eigenvectors of $A$
   II. $A$ is diagonalizable
   III. Rows of $A$ are linearly independent

   **(A)** I
   **(B)** II
   **(C)** I and II
   **(D)** II and III
   **(E)** I and III

4. Which of the following could be the set of eigenvalues of $\begin{pmatrix} 5 & -3 & 3 \\ 3 & -5 & 3 \\ -3 & -1 & -1 \end{pmatrix}$?

**(A)** $\lambda_1 = 2$, $\lambda_2 = 1$, $\lambda_3 = 2$
**(B)** $\lambda_1 = 2$, $\lambda_2 = -1$, $\lambda_3 = 2$
**(C)** $\lambda_1 = 2$, $\lambda_2 = 1$, $\lambda_3 = -2$
**(D)** $\lambda_1 = 2$, $\lambda_2 = -1$, $\lambda_3 = -2$
**(E)** None of these

5. The quadratic form $8x_1^2 - 2x_1x_2 + 2x_1x_3 + x_2^2 - 2x_2x_3 - 7x_3^2$ is

**(A)** positively definite
**(B)** negatively definite
**(C)** non-positively definite
**(D)** non-negatively definite
**(E)** not definite

6. Let $d$ be the distance between the endpoint of vector $x$ and the linear space generated by vectors $e_1$, and $e_2$, where $x = \begin{pmatrix} -2 \\ -7 \\ -11 \end{pmatrix}$, $e_1 = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}$, $e_2 = \begin{pmatrix} -2 \\ -1 \\ -1 \end{pmatrix}$ Then, $d^2 =$

**(A)** 2
**(B)** 4
**(C)** 8
**(D)** 16
**(E)** None of these

7. Which of the following is NOT an eigenvalue of $A = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 5 & -5 & -5 & -4 \\ 0 & 0 & 0 & -3 \end{pmatrix}$?

**(A)** -5
**(B)** -3
**(C)** 0
**(D)** 3
**(E)** 5

8. Let $A$ be a 4×4 matrix with $\det(A) \neq 0$. Which of the following MIGHT be true?

I. $A$ has no real eigenvalues

II. $\lambda = 0$ is an eigenvalue of $A$

III. $A$ and $A^2$ have the same set of eigenvalues

**(A)** I
**(B)** II
**(C)** III
**(D)** I and II
**(E)** I and III

9. The element $g_{32}$ of the Gram matrix of $e_1 = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$, $e_3 = \begin{pmatrix} -3 \\ 4 \\ -1 \end{pmatrix}$ is

**(A)** -11
**(B)** 1
**(C)** 6
**(D)** 13
**(E)** None of these

10. Let $A$ be a 3×3 matrix whose characteristic polynomial $f(\lambda) = 1 + \lambda - \lambda^2 - \lambda^3$. Which of the following MUST be true?

I rk$(A - \lambda E) = 3$ for some $\lambda$

II. $A^{-1}$ exists

III. $A$ is NOT diagonalizable

**(A)** I
**(B)** II
**(C)** I and II
**(D)** II and III
**(E)** I, II, and III

## SECTION II: FREE RESPONSE

**Directions:** Show all your work. Indicate clearly the methods you use because you will be graded on the correctness of your methods as well as on the accuracy of your results and explanations. *You may use the back of this sheet for writing.*

1. [**6 pts**] For the following matrix $A$ find matrices $C$ and $C^{-1}$ such that $C^{-1}AC$ is diagonal and write a product expression for $A^{100}$. Show your work.

$$A = \begin{pmatrix} -1 & 5 & 3 \\ -3 & 3 & 3 \\ 3 & 1 & -1 \end{pmatrix}$$

2. [**4 pts**] Use Gram-Schmidt orthogonalization process to find an orthogonal base in the linear hull of the following system of vectors. Show your work.

$$e_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 3 \end{pmatrix}, e_2 = \begin{pmatrix} -3 \\ 0 \\ -3 \\ 1 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 3 \\ -3 \\ 1 \end{pmatrix}$$

3. [**4 pts**] For the vectors $x$, $e_1$, $e_2$, and $e_3$ below, find the projection of the vector $x$ onto the linear hull of $e_1$, $e_2$, and $e_3$. Evaluate the distance between the endpoint of $x$ and the linear hull.

$$x = \begin{pmatrix} 3 \\ 2 \\ 2 \\ 1 \end{pmatrix}, e_1 = \begin{pmatrix} -4 \\ 4 \\ 4 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 2 \\ 2 \\ 1 \\ -4 \end{pmatrix}, e_3 = \begin{pmatrix} -3 \\ 2 \\ 3 \\ 1 \end{pmatrix}$$

# ANSWERS

**Multiple Choice:**

1. B
2. D
3. C
4. D
5. E
6. C
7. E
8. E
9. A
10. C

**Free Response:**

1. $\begin{pmatrix} -3 & 4 & 3 \\ 1 & -3 & -2 \\ -1 & 1 & 1 \end{pmatrix} \times \begin{pmatrix} -1 & 5 & 3 \\ -3 & 3 & 3 \\ 3 & 1 & -1 \end{pmatrix} \times \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & -3 \\ -2 & -1 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

2. $f_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 3 \end{pmatrix}$, $f_2 = \begin{pmatrix} -7 \\ 1 \\ -7 \\ 4 \end{pmatrix}$, $f_3 = \begin{pmatrix} 10 \\ 15 \\ -13 \\ -9 \end{pmatrix}$

3. $p = \begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \end{pmatrix}$, $d^2 = 12$

# 15 Sample Final Exam III

## SECTION I: MULTIPLE CHOICE

**Directions:** Each of the following problems is followed by five choices. Select the best choice and put the corresponding mark in your answer sheet. **Calculators may NOT be used at any part of the exam.**

1. The quadratic form $x_1^2 - 2x_1x_2 - 4x_1x_3 + 4x_2^2 + 4x_2x_3 + 4x_3^2$ is
   **(A)** positive semi-definite
   **(B)** positive definite
   **(C)** negative semi-definite
   **(D)** negative definite
   **(E)** not definite

2. Let A and B be n×n matrices such that $B \cdot A \cdot B = E$. Which of the following COULD be true?

   I. $\det(A) \leq 0$

   II. $\mathrm{rk}A < \mathrm{rk}B$

   III. $A \cdot B \neq B \cdot A$

   **(A)** I only
   **(B)** II only
   **(C)** III only
   **(D)** II and III
   **(E)** None could be true

3. If $d$ is the distance between the endpoint of vector $x = (-2, 7, -1)$ and the linear hull of vectors $e_1 = (1, 0, 0)$ and $e_2 = (1, -3, -3)$, then $d^2 =$
   **(A)** 4
   **(B)** 16
   **(C)** 32
   **(D)** 48
   **(E)** None of these

4. Let $A : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear operator and let $e_1$, $e_2$, $e_3$ be some of its eigenvectors. Given that $e_1$, $e_2$, $e_3$ are linearly independent, which of the following MUST be true?

   I. $A$ is diagonalizable

   II. The eigenvalues corresponding to $e_1$, $e_2$, and $e_3$ are different

   III. $A^{-1}$ is diagonalizable, if exists

**(A)** I only
**(B)** I and II
**(C)** I and III
**(D)** I, II, and III
**(E)** None

5. Let $A$ be a n×n matrix and let $B$ be the matrix defined by $B = A^T \cdot A$. Which of the following are true statements?

   I. Regardless of $A$, matrix $B$ is always symmetric

   II. $B$ is the matrix of a positive semi-definite quadratic form

   III. $B$ is the matrix of a positive definite quadratic form if and only if $\det(A) \neq 0$

   **(A)** I only
   **(B)** I and II
   **(C)** I and III
   **(D)** II and III
   **(E)** I, II, and III

6. Let $\mathcal{A}$ be the linear operator acting on infinite-differentiable real functions according to the rule $\mathcal{A}(f) = \frac{\partial^2 f}{\partial x^2} + f$. Which of the following is an eigenvector of $\mathcal{A}$?
   **(A)** $f(x) = \cos 2x$
   **(B)** $f(x) = \sin 3x$
   **(C)** $f(x) = e^{-4x}$
   **(D)** All of the above
   **(E)** None of the above

7. Which of the following matrix equations, where $X$ and $Y$ are the unknown matrices, MUST have no solutions?

   I. $X^2 = -E$

   II. $X^T \cdot X = -E$

   III. $X \cdot Y - Y \cdot X = E$

   **(A)** I only
   **(B)** I and III
   **(C)** III only
   **(D)** II and III
   **(E)** I, II, and III only

8. Which of the following numbers is NOT an eigenvalue of the matrix $A = \begin{pmatrix} 3 & -7 & -7 \\ 7 & -5 & -7 \\ 1 & -7 & -5 \end{pmatrix}$?

   **(A)** -5

**(B)** 2
**(C)** 7
**(D)** -4
**(E)** All are eigenvalues

9. The SLE $\begin{cases} ax_1 - x_2 + 2x_3 = 11 \\ 2x_1 + 9x_2 - 4x_3 = 0 \\ 2x_1 - 5x_2 + 3x_3 = -7 \end{cases}$ has no solutions when **(A)** $a = -5$

**(B)** $a = 6$
**(C)** $a = 0$
**(D)** $a = 5$
**(E)** There is no such $a$

10. Let $e_1$, $e_2$, and $e_3$ be three distinct eigenvectors of a linear operator $A : \mathbb{R}^3 \to \mathbb{R}^3$. If $e_1$, $e_2$, and $e_3$ are linearly dependent, which of the following MUST be true?

I. $\det(A) = 0$

II. At least two of the eigenvalues corresponding to $e_1$, $e_2$, $e_3$ must be equal to each other

III. $A$ is not diagonalizable

**(A)** I only
**(B)** II only
**(C)** II and III
**(D)** I, II, and III
**(E)** None

## SECTION II: FREE RESPONSE

**Directions:** Show all your work. Indicate clearly the methods you use because you will be graded on the correctness of your methods as well as on the accuracy of your results and explanations. *You may use the back of this sheet for writing.*

1. **[3 pts]** Find the fundamental set of solutions to the following system of linear equations. Indicate the dimension of the solution space.

$$\begin{cases} x_1 + 2x_2 + 5x_3 + 5x_4 + 8x_5 = -2 \\ x_1 + x_2 + 6x_3 + 4x_4 = -5 \end{cases}$$

2. **[4 pts]** For the following matrix $A$ find matrices $C$ and $C^{-1}$ such that $C^{-1}AC$ is diagonal and write a product expression for $A^n$. Show your work.

$$\begin{pmatrix} 5 & 1 & -2 \\ -2 & 2 & 4 \\ -1 & -1 & 6 \end{pmatrix}$$

3. **[3 pts]** Use Gram-Schmidt orthogonalization process to find an orthogonal base in the linear hull of the following system of vectors. Show your work.

$$e_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, e_2 = \begin{pmatrix} 1 \\ 1 \\ 4 \\ 2 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -3 \end{pmatrix}.$$

# ANSWERS

**Multiple Choice:**

1. A
2. E
3. C
4. C
5. E
6. D
7. D
8. C
9. B
10. B

**Free Response:**

1. $x = \begin{pmatrix} -8 \\ 3 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -7 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -3 \\ -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 8 \\ -8 \\ 0 \\ 0 \\ 1 \end{pmatrix}$,
   dim = 3.

2. $\begin{pmatrix} 2 & 3 & -4 \\ 1 & 1 & -1 \\ -1 & -1 & 2 \end{pmatrix} \times \begin{pmatrix} 5 & 1 & -2 \\ -2 & 2 & 4 \\ -1 & -1 & 6 \end{pmatrix} \times \begin{pmatrix} -1 & 2 & -1 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix}$,
   which is also correct.

3. $f_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $f_2 = \begin{pmatrix} -3 \\ 4 \\ -5 \\ 1 \end{pmatrix}$, $f_3 = \begin{pmatrix} -8 \\ 22 \\ 15 \\ -37 \end{pmatrix}$
   Note: some answers are not unique.

# Index