Linear Algebra with Sage (BigBook, Free e-book, English Version) All

Book · October 2015

8 authors, including:

Sang-Gu Lee
Sungkyunkwan University

In-Jae Kim
Minnesota State University, Mankato

Phong Vu
Vietnam National University, Hanoi

Victoria Lang
Sungkyunkwan University

ISBN 978-89-24-01524-9

V.11

July 7, 2015

Linear Algebra

Translated by

Sang-Gu LEE with Jon-Lark KIM, In-Jae KIM, Namyong LEE,
Ajit KUMAR, Phong VU, Victoria LANG, Jae Hwa LEE

(Based on the book written by Sang-Gu Lee with Jae Hwa Lee, Kyung-Won Kim)
http://matrix.skku.ac.kr/2015-Album/BigBook-LinearAlgebra-2015.pdf

http://sage.skku.edu, http://www.sagemath.org
and http://matrix.skku.ac.kr/LA-Lab/

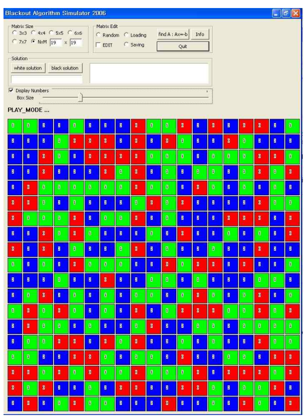
BIGBOOK
공유와 협력의 교과서 만들기 운동본부

http://www.bigbook.or.kr/

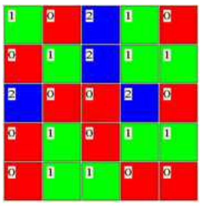
Linear Algebra with

http://matrix.skku.ac.kr/LA-Sage/

Contents



Preface



This book, 'Linear Algebra with Sage', has two goals. The first goal is to explain Linear Algebra with the help of Sage. Sage is one of the most popular computer algebra system(CAS). Sage is a free and user-friendly software. Whenever the Sage codes are possible, we illustrate examples with Sage codes. The second goal is to make the book accessible to everyone in the world freely. Therefore, the pdf file of this book is free to use in class or in person. For commercial use, please contact us.

Linear Algebra is regarded as one of the most important mathematical subjects because it is used not only in natural sciences and engineering applications but also in humanities and social sciences. Nowadays, Linear Algebra is studied most actively in the 21st century.

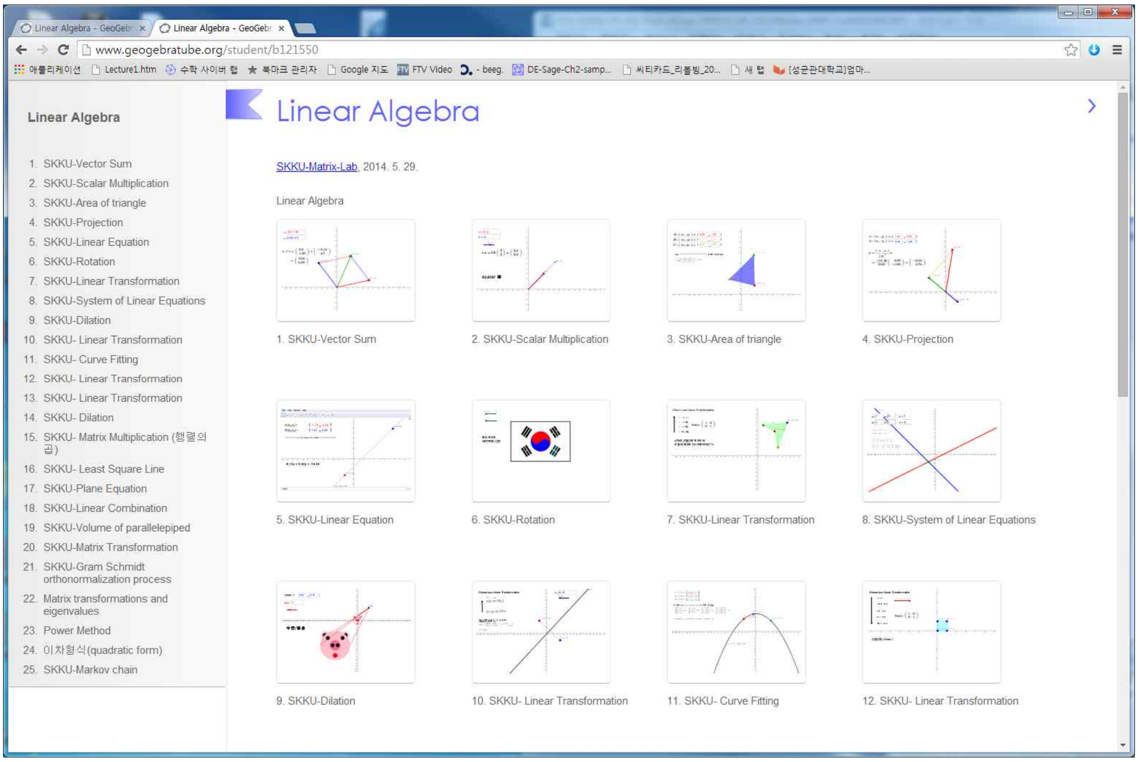
One of the roles of mathematics in society is to suggest a possible solution by modeling a practical problem as a mathematical problem, by solving it with the idea of a system of linear equations, and by interpreting the solution in the setting of the original problem. The first computer is also based on the linear process. The study and applications of Linear Algebra grew incredibly in the later part of the 20th century.

It is interesting to note that Sylvester and Cayley, inventors of matrices, and Babbage, father of the computer, were mathematicians in the 19th century from United Kingdom. Since then, the study of matrix theory has progressed and contributed to the development of physics by the appearance of infinite dimensions and tensors.

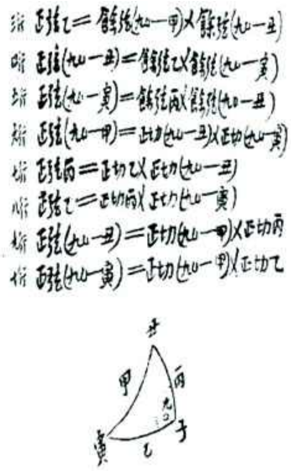
Matrix theory in the United States of America was neglected from the European mathematical society before the Second World War. After that, because the modern computers were built and the numerical power of matrices became very useful, the matrix theory was developed well in the United Sates in the 20th century. The United States has grown as a unique super power in both theories and experiments of sciences.

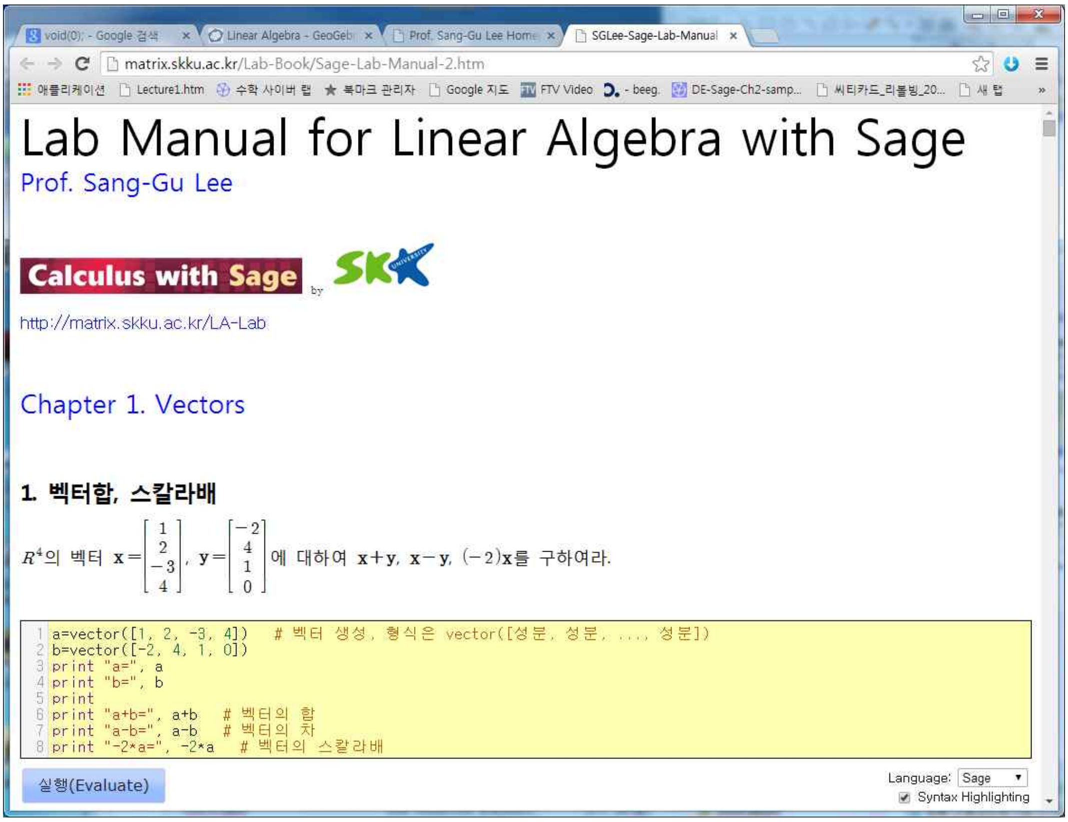
How to use Lab

https://www.youtube.com/watch?v=V0xJvW-YjWs

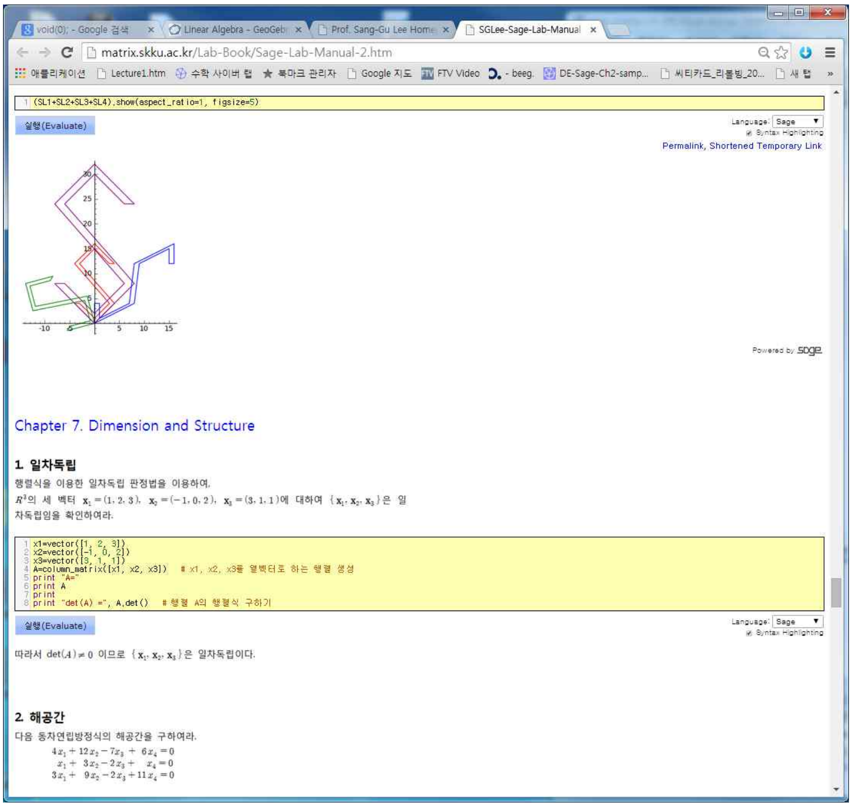


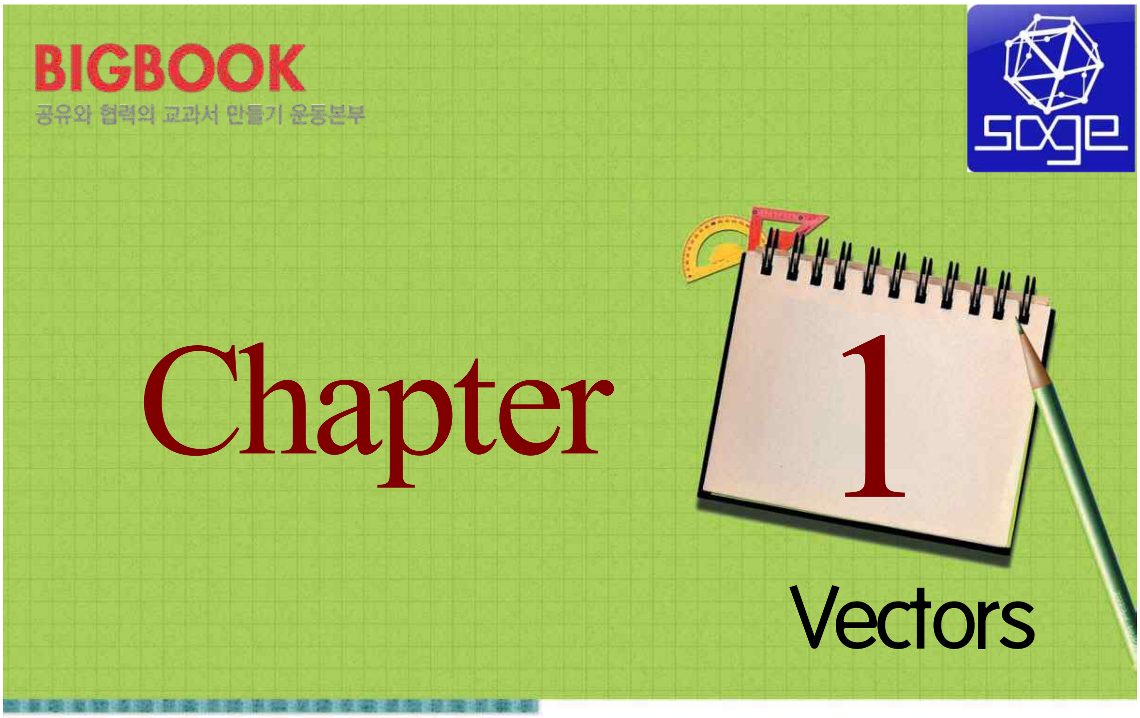
[CAS-Geogebra] http://www.geogebratube.org/student/b121550





[CAS-Sage] http://matrix.skku.ac.kr/knou-knowls/Sag-Ref.htm





Chapter 1 Vectors

Linear algebra is the branch of mathematics concerning vectors and mappings. Linear algebra is central to both pure and applied mathematics. Combined with calculus, linear algebra facilitates the solution of linear systems of differential equations. Techniques from linear algebra are also used in analytic geometry, engineering, physics, natural sciences, computer science, computer animation, and the social sciences (particularly in economics). A geometric quantity described by a magnitude and a direction is called a vector. In this chapter, we begin with studying basic properties of vectors starting from 3-dimensional vectors and extending these properties to n-dimensional vectors. We will also discuss the notion of the dot product (or inner product) of vectors and vector equations of lines and planes.

Introduction : http://youtu.be/Mxp1e2Zzg-A

1.1 *Vectors in n-space

- Reference video: http://youtu.be/aeLVQoPQMpE http://youtu.be/85kGK6bJLns
- Practice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-1-Sec-1-1.html

Among the physical quantities we use and encounter in everyday life, scalar (e.g. length, area, mass, temperature, etc.) is a quantity that can be completely described by a single real number. A vector (e.g. force, velocity, change in position, etc.) is a geometric quantity described by a magnitude and a direction.

- Scalar: length, area, mass, temperature- a one-dimensional physical quantity, i.e. one that can be described by a single real number.

- Vector : velocity, change in position, force - a geometric quantity described by a magnitude and a direction.

- A vector can be sketched as a directed line segment; in 2-and 3-dimensional space, vectors are often drawn as arrows.

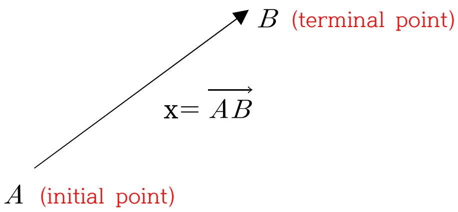


Figure 1

- A vector with the same initial and terminal points with magnitude 0 is called the **zero (or null) vector.** (Since its magnitude is 0, it does not have a specific direction).

- In physics, vectors provide a useful way to express velocity, acceleration, force, and the laws of motion. A force vector can be broken down into mutually perpendicular component vectors. An **electric field** can be visualized by field vectors, which indicate both the magnitude and direction of the field at every point in space. Vectors have a wide variety of applications in the social sciences, such as population dynamics and economics.

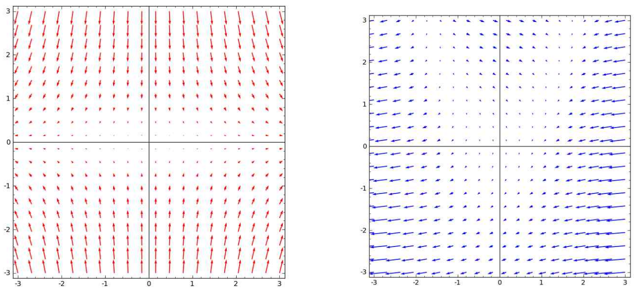


Figure 2

● From now on, unless noted otherwise, we will restrict scalars to real numbers - that is, if k is a scalar, $k \in \mathbb{R}$.

Definition **[Vector Addition and Scalar Multiplication]**

For any two vectors $\mathbf{x}$, $\mathbf{y}$, and scalar k, the **sum** of $\mathbf{x}$ and $\mathbf{y}$, $\mathbf{x}+\mathbf{y}$, and the **scalar multiple** of $\mathbf{x}$ by k, $k\mathbf{x}$, are defined as follows.

(1) The sum of $\mathbf{x}$ and $\mathbf{y}$ is found by placing $\mathbf{x}$ and $\mathbf{y}$ tail-to-tail to form two adjacent sides of a parallelogram. The diagonal of this parallelogram is $\mathbf{x}+\mathbf{y}$. This is called the *Parallelogram Law*. (See Figure 3.)

(2) The **scalar multiple** of $\mathbf{x}$ by a scalar k, is a vector with magntitude $|k|$ times the magnitude of $\mathbf{x}$ and with the same direction as $\mathbf{x}$ if $k>0$, and is opposite to $\mathbf{x}$ if $k<0$. (See Figure 4.) If k is 0, $k\mathbf{x}$ is the zero vector.

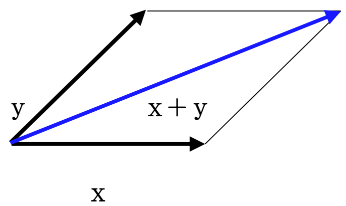


Figure 3

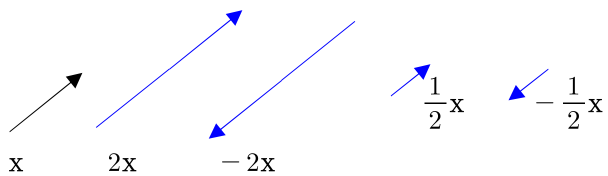


Figure 4

In the real coordinate plane $\mathbb{R}^2 = \{(x_1, x_2) \mid x_1, x_2 \in \mathbb{R}\}$, the initial and terminal points of every vector determine its the magnitude and direction. If vectors have the same magnitude and direction, even if they are in different positions, we regard these vectors as **equivalent**.

Definition

An ordered pair of real numbers (x_1, x_2) is called a **vector (in $\mathbb{R}^2$)** and can be written as

$$\mathbf{x} = (x_1, x_2) \text{ or } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Here, x_1, x_2 are called the **components** of $\mathbf{x}$.

Definition **Equivalence**

Two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ with $x_1 = y_1$, $x_2 = y_2$, then we say that $\mathbf{x}$ and $\mathbf{y}$ are **equivalent** (or **equal**) and we write $\mathbf{x} = \mathbf{y}$.

[Remark] ***The case when the initial point is not at the origin.***

A directed line from the point $P(x_1, x_2)$ to the point $Q(y_1, y_2)$ is a vector with the following components: $\overrightarrow{PQ} = \overrightarrow{OQ'} = (y_1 - x_1, y_2 - x_2)$. The initial point of the vector $\overrightarrow{OQ} = (x_1, x_2)$ is at the origin $O(0, 0)$ and the terminal point is $P(x_1, x_2)$.

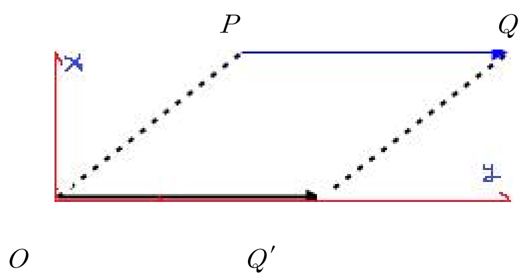


Figure 5

Example 1

For $O(0,0)$, $P_1(0, -4)$, $P_2(-3, 1)$, $Q(2, 3)$, $Q_1(2, -1)$, $Q_2(-1, 4) \in \mathbb{R}^2$, express the vectors $\overrightarrow{OQ}$, $\overrightarrow{P_1Q_1}$, $\overrightarrow{P_2Q_2}$ in component form.

Solution $\overrightarrow{OQ} = (2, 3)$, $\overrightarrow{P_1Q_1} = \overrightarrow{OQ_1} - \overrightarrow{OP_1} = (2, -1) - (0, -4) = (2, 3)$,
$\overrightarrow{P_2Q_2} = \overrightarrow{OQ_2} - \overrightarrow{OP_2} = (-1, 4) - (-3, 1) = (2, 3)$

$\overrightarrow{P_1Q_1}$ and $\overrightarrow{P_2Q_2}$ are equivalent.

Sage sol. Copy the following code into http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/ to practice.

```
o=vector([0, 0])      #creates a vector, x=vector([component x1, component x2])
p1=vector([0, -4])
p2=vector([-3, 1])
q=vector([2, 3])
q1=vector([2, -1])
q2=vector([-1, 4])
print "vector OQ=", q-o               # subtract
print "vector P1Q1=", q1-p1           # subtract
print "vector P2Q2=", q2-p2           # subtract
print "vector OQ = vector P1Q1= vector P2Q2"
```

```
vector OQ= (2, 3)
vector P1Q1= (2, 3)
vector P2Q2= (2, 3)
vector vector OQ = P1Q1= vector P2Q2
```

■

Definition

For any two vectors $\mathbf{x}=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\mathbf{y}=\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ in $\mathbb{R}^2$ and scalar k, the **sum** of $\mathbf{x}$ and $\mathbf{y}$, $\mathbf{x}+\mathbf{y}$, and the **scalar multiple** of $\mathbf{x}$ by k, $k\mathbf{x}$, are defined *component-wise* as follows.

$$\text{(i)}\ \ \mathbf{x}+\mathbf{y}=\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \end{bmatrix} \qquad\qquad \text{(ii)}\ \ k\mathbf{x}=\begin{bmatrix} kx_1 \\ kx_2 \end{bmatrix}$$

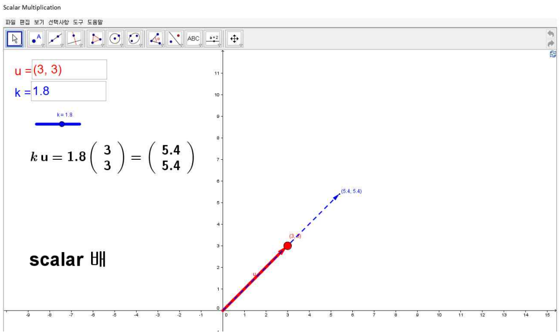
In $\mathbb{R}^2$, the **zero vector** is a vector where all its components are equal to 0 (its initial point is taken to be the origin). Then, for an arbitrary $\mathbf{x}$ in R^2, it is clear that

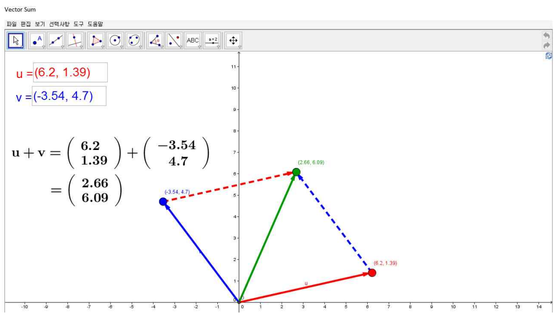
$$\mathbf{x}+\mathbf{0}=\mathbf{x},\ \ \mathbf{x}+(-1)\mathbf{x}=\mathbf{0}.$$

Here, taking $(-1)\mathbf{x}=-\mathbf{x}$, we call $-\mathbf{x}$ the **negative vector** or additive inverse of $\mathbf{x}$.

[Remark] **Computer Simulations**

[Scalar multiplication] http://matrix.skku.ac.kr/2012-album/2.html
[Vector addition] http://matrix.skku.ac.kr/2012-album/3.html





Example **2**

For vectors $\mathbf{x} = \begin{bmatrix}1\\2\end{bmatrix}$, $\mathbf{y} = \begin{bmatrix}-2\\4\end{bmatrix}$ in $\mathbb{R}^2$, find $\mathbf{x}+\mathbf{y}$, $\mathbf{x}-\mathbf{y}$, and $(-2)\mathbf{x}$.

- http://matrix.skku.ac.kr/RPG_English/1-VT-sum-multi.html

Sage sol. Copy the following code into http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/ to practice.

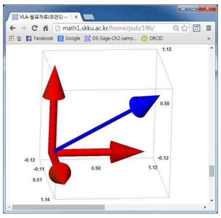
```
x=vector([1, 2])      #creates a vector, x=vector([component x_1, component x_2])
y=vector([-2, 4])
print "x+y=", x+y              # adds vectors
print "x-y=", x-y              # subtract vectors
print "-2*x=", -2*x            # multiplies vectors by scalar, (you must
                                 include '*' when multiplying)
```

```
x+y=(-1, 6)
x-y=(3, -2)
-2*x=(-2, -4)
```
■

- In $\mathbb{R}^3 = \{(x_1, x_2, x_3) \mid x_1, x_2, x_3 \in \mathbb{R}\}$, we define vectors as follows.

Definition

A 3-tuple of real numbers (x_1, x_2, x_3) is called a **vector (in $\mathbb{R}^3$)** and can be written as



$$\mathbf{x} = (x_1, x_2, x_3) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{3\times 1}.$$

Here, x_1, x_2, x_3 are called the **components** of $\mathbf{x}$.

Definition [Equivalence or Equality]

Two vectors $\mathbf{x}, \mathbf{y} \in R^3$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ with $x_1 = y_1$, $x_2 = y_2$, $x_3 = y_3$, are said to be **equivalent** (or **equal**) and we write $\mathbf{x} = \mathbf{y}$.

[Remark] *The case when the initial point is the origin.*

A directed line from the pont $P(x_1, x_2, x_3)$ to the point $Q(y_1, y_2, y_3)$ is a vector with the following components: $\overrightarrow{PQ} = (y_1 - x_1, y_2 - x_2, y_3 - x_3) = \overrightarrow{OQ'}$.

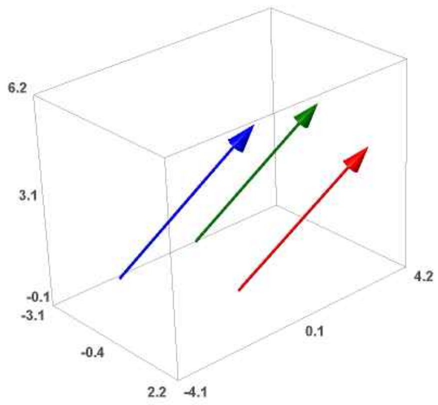


Figure 6

Example 3

For $P_1(0,-4,2)$, $P_2(-3,1,0)$, $Q(2,3,4)$, $Q_1(2,-1,6)$, $Q_2(-1,4,4) \in \mathbb{R}^3$, express the vectors $\overrightarrow{OQ}$, $\overrightarrow{P_1Q_1}$, $\overrightarrow{P_2Q_2}$ in component form.

Solution

$\overrightarrow{OQ} = (2,3,4)$, $\overrightarrow{P_1Q_1} = \overrightarrow{OQ_1} - \overrightarrow{OP_1} = (2,-1,6)-(0,-4,2) = (2,3,4)$,

$\overrightarrow{P_2Q_2} = \overrightarrow{OQ_2} - \overrightarrow{OP_2} = (-1,4,4)-(-3,1,0) = (2,3,4)$

$\overrightarrow{P_1Q_1}$ and $\overrightarrow{P_2Q_2}$ are equivalent.

Sage sol. Copy the following code into http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/ to practice.

```
o=vector([0, 0, 0])     #creates a vector
p1=vector([0, -4, 2])
p2=vector([-3, 1, 0])
q=vector([2, 3, 4])
q1=vector([2, -1, 6])
q2=vector([-1, 4, 4])
print "vector OQ=", q-o              # subtract
print "vector P1Q1=", q1-p1          # subtract
print "vector P2Q2=", q2-p2          # subtract
print "vector OQ = vector P1Q1= vector P2Q2"
```

```
vector OQ= (2, 3, 4)
vector P1Q1= (2, 3, 4)
vector P2Q2= (2, 3, 4)
vector OQ = vector P1Q1= vector P2Q2
```

■

Definition

For any two vectors

$$\mathbf{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix},\quad \mathbf{y}=\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$$

in $\mathbb{R}^3$ and scalar k, the **sum** of $\mathbf{x}$ and $\mathbf{y}$, $\mathbf{x}+\mathbf{y}$, and the **scalar multiple** of $\mathbf{x}$ by k, $k\mathbf{x}$, are defined *component-wise* as follows:

$$\text{(i)}\ \ \mathbf{x}+\mathbf{y}=\begin{bmatrix}x_1+y_1\\x_2+y_2\\x_3+y_3\end{bmatrix} \qquad\qquad \text{(ii)}\ \ k\mathbf{x}=\begin{bmatrix}kx_1\\kx_2\\kx_3\end{bmatrix}.$$

In $\mathbb{R}^3$, the **zero vector** is a vector where all its components are equal to 0 (its initial point is taken to be the origin). Then, for an arbitrary $\mathbf{x}$ in $\mathbb{R}^3$, it is clear that

$$\mathbf{x}+\mathbf{0}=\mathbf{x},\ \ \mathbf{x}+(-1)\mathbf{x}=\mathbf{0}.$$

Here, taking $(-1)\mathbf{x}=-\mathbf{x}$, we call $-\mathbf{x}$, the **negative vector** of $\mathbf{x}$.

- The **Euclidean spaces** $\mathbb{R}^2$ and $\mathbb{R}^3$ can be generalized to **n-dimensional Euclidean space** $\mathbb{R}^n$ as follows:

$$\mathbb{R}^n=\{(x_1,x_2,\dots,x_n)\mid x_i\in\mathbb{R},\ i=1,2,\dots,n\}$$

$\mathbb{R}^n$ is also called n-dimensional space and elements of $\mathbb{R}^n$ are called n-dimensional vectors. (We shall formally define vector space later.)

Definition

An ordered n-tuple of real numbers $(x_1,x_2,\dots,x_n)$ is called a **n-dimensional vector** and can be written as

$$\mathbf{x}=(x_1,x_2,\dots,x_n)\ \ =\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}_{n\times 1}$$

Here, real numbers $x_1,x_2,\dots,x_n$ are called the **components** of $\mathbf{x}$.

Definition **[Equivalence or Equality]**

For vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

if $x_i = y_i$ $(i = 1, 2, \ldots, n)$ then we say $\mathbf{x}$ and $\mathbf{y}$ are **equivalent** (or **equal**) and we write $\mathbf{x} = \mathbf{y}$.

Definition

For any two vectors

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

in $\mathbb{R}^n$ and scalar k, the **sum** of $\mathbf{x}$ and $\mathbf{y}$, $\mathbf{x}+\mathbf{y}$, and the **scalar multiple** of $\mathbf{x}$ by k, $k\mathbf{x}$, are defined *component-wise* as follows:

$$\text{(i)} \ \ \mathbf{x}+\mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix} \qquad \text{(ii)} \ \ k\mathbf{x} = \begin{bmatrix} kx_1 \\ kx_2 \\ \vdots \\ kx_n \end{bmatrix}.$$

In $\mathbb{R}^n$, the **zero vector** is a vector where all its components are equal to 0 (its initial point is taken to be the origin). Then, for an arbitrary $\mathbf{x}$ in $\mathbb{R}^n$, it is clear that

$$\mathbf{x} + \mathbf{0} = \mathbf{x}, \ \ \mathbf{x} + (-1)\mathbf{x} = \mathbf{0}.$$

Here, taking $(-1)\mathbf{x} = -\mathbf{x}$, we call $-\mathbf{x}$, the **negative vector** of $\mathbf{x}$.

Example 4

Find $\mathrm{x}+\mathrm{y}$, $\mathrm{x}-\mathrm{y}$, $(-2)\mathrm{x}$ when $\mathrm{x}=\begin{bmatrix}1\\2\\-3\\4\end{bmatrix}$ and $\mathrm{y}=\begin{bmatrix}-2\\4\\1\\0\end{bmatrix}$ in $\mathbb{R}^4$.

- http://matrix.skku.ac.kr/RPG_English/1-VT-sum-multi-3.html

Solution

$\mathrm{x}+\mathrm{y}=(1,2,-3,4)+(-2,4,1,0)=(-1,6,-2,4)$,
$\mathrm{x}-\mathrm{y}=(1,2,-3,4)-(-2,4,1,0)=(3,-2,-4,4)$,
$(-2)\mathrm{x}=(-2)(1,2,-3,4)=(-2,-4,6,-8)$ □

Sage Copy the following code into http://sage.skku.edu to practice.

```
x=vector([1, 2, -3, 4])
y=vector([-2, 4, 1, 0])
print "x+y=", x+y                # adds vectors
print "x-y=", x-y                # subtracts vectors
print "-2*x=", -2*x
```

```
x+y=(-1, 6, -2, 4)
x-y=(3, -2, -4, 4)
-2*x=(-2, -4, 6, -8)
```

■

Theorem 1.1.1

If x, y, z are vectors in $\in \mathbb{R}^n$ and h and k are scalars, then

(1) $\mathrm{x}+\mathrm{y}=\mathrm{y}+\mathrm{x}$
(2) $(\mathrm{x}+\mathrm{y})+\mathrm{z}=\mathrm{x}+(\mathrm{y}+\mathrm{z})$
(3) $\mathrm{x}+0=\mathrm{x}=0+\mathrm{x}$
(4) $\mathrm{x}+(-\mathrm{x})=0=(-\mathrm{x})+\mathrm{x}$
(5) $k(\mathrm{x}+\mathrm{y})=k\mathrm{x}+k\mathrm{y}$
(6) $(h+k)\mathrm{x}=h\mathrm{x}+k\mathrm{x}$
(7) $(hk)\mathrm{x}=h(k\mathrm{x})$
(8) $1\mathrm{x}=\mathrm{x}$

The proof of above theorem is simple and follows from properties of addition and multiplication of real numbers.

Theorem 1.1.2

If $\mathbf{x}$ is a vector in $\in \mathbb{R}^n$ and k is a scalar, then

(1) $0\,\mathbf{x} = \mathbf{0}$
(2) $k\,\mathbf{0} = \mathbf{0}$
(3) $(-1)\,\mathbf{x} = -\mathbf{x}$

Definition

For vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in $\mathbb{R}^n$ and scalars $c_1, c_2, \ldots, c_k$,

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

is called a **linear combination** of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$.

Example 5

Find $2\mathbf{x} - 3\mathbf{y} + \mathbf{z}$, when $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -2 \\ 4 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{z} = \begin{bmatrix} 5 \\ -2 \\ 3 \\ 7 \end{bmatrix}$ in $\mathbb{R}^4$.

- http://matrix.skku.ac.kr/RPG_English/1-VT-sum-multi-3.html

Solution

$2\mathbf{x} - 3\mathbf{y} + \mathbf{z} = (13, -10, -6, 1)$

Sage Copy the following code into http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/ to practice.

```
x=vector([1, 2, -3, 4])
y=vector([-2, 4, 1, 0])
z=vector([5, -2, 3, -7])
print "2*x-3*y+z=", 2*x-3*y+z                # linear combination
```

```
2*x-3*y+z=(13, -10, -6, 1)
```

■

Example 6 The above **Example 5** can also be done in Sage as follows. First, we build the relevant vectors and the command for a linear combination of many vectors. Then, we can combine all into one line, as follows.

Sage Copy the following code into http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/ to practice.

```
x=vector(QQ, [1, 2, -3, 4])  # computations with quotient numbers in Q
y=vector(QQ, [-2, 4, 1, 0])
z=vector(QQ, [5, -2, 3, -7])
print "2*x-3*y+z=", 2*x-3*y+z                # linear combination
vectors = [x, y, z]
scalars = [2, -3, 1]
multiples = [scalars[i]*vectors[i] for i in range(3)]
print "a*x+b*y+c*z=", sum(multiples)          # linear combination
```

```
2*x-3*y+z = (13, -10, -6, 1)
a*x+b*y+c*z = (13, -10, -6, 1)
```
■

(Comment : We can create an applet to generate a random vectors and scalars and find the linear combination, as well.**)**

Rob Beezer's Linear Combination Lab: http://linear.ups.edu/html/section-LC.html

<Sang-Seol LEE, Father of Korean Mathematics education>

http://www.youtube.com/watch?feature=player_embedded&v=NbuRcvLlJOw


Courrier de la Conférence

DE LA PAIX

Rédigé par WILLIAM T. STEAD

Publié sous les auspices de la FONDATION POUR L'INTERNATIONALISME à La Haye

Directeurs Editeurs: MAAS & VAN SUCHTELEN

REAUX: Princessegracht 6A, La Haye. — Telephone No. 287 — Adr. Télégr. MAASSUCHTELEN

AVIS

Qu'est-ce que la Contrebande de Guerre?

1.2 Inner product and Orthogonality

- Reference videos: http://youtu.be/g55dfkmlTHE , http://youtu.be/CbfJYPCkbm8
- Practice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-1-Sec-1-1.html

In this section, we will discuss the concepts of vector length, distance, and how to calculate the angle between two vectors, as well as vector parallelism and orthogonality in $\mathbb{R}^n$.

Definition

Given a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ in $\mathbb{R}^n$

$$\| \mathbf{x} \| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

is called the **norm (or length or magnitude)** of $\mathbf{x}$, and is denoted by the symbol $|\mathbf{x}|$ or $\|\mathbf{x}\|$ (read as norm $\mathbf{x}$).

In the above definition, $\| \mathbf{x} \|$ is the distance from the initial point of the vector $\mathbf{x}$ to its terminal point; equivalently, it is the distance from the origin to the point $P(x_1, x_2, \ldots, x_n)$. Therefore, for any two vectors $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ in R^n, $\| \mathbf{x} - \mathbf{y} \|$ is the distance between the two points $P(x_1, x_2, \ldots, x_n)$ and $Q(y_1, y_2, \ldots, y_n)$. That is,

$$\| \mathbf{x} - \mathbf{y} \| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}\ .$$

Example 1

For the vectors $\mathbf{x} = (2, -1, 3, 2)$, $\mathbf{y} = (3, 2, 1, -4)$ in R^4, we have the following.

- http://matrix.skku.ac.kr/RPG_English/1-B1-norm-distance.html

Solution

$$\| \mathbf{x} \| = \sqrt{2^2 + (-1)^2 + 3^2 + 2^2} = \sqrt{4+1+9+4} = 3\sqrt{2}$$

$$\begin{aligned} \| \mathbf{y} \| &= \sqrt{3^2 + 2^2 + 1^2 + (-4)^2} \\ &= \sqrt{9+4+1+16} = \sqrt{30} \end{aligned}$$

$$\| \mathbf{x}-\mathbf{y} \| = \sqrt{(2-3)^2+(-1-2)^2+(3-1)^2+(2-(-4))^2} = \sqrt{50}$$
$$= 5\sqrt{2} \ . \qquad \square$$

Sage Copy the following code into http://sage.skku.edu to practice.

```
x=vector([2, -1, 3, 2])
y=vector([3, 2, 1, -4])
print x.norm()              # calculate the norm of x
print y.norm()              # calculate the norm of y
print (x-y).norm()          # calculate distance
```

```
3*sqrt(2)                   # sqrt(2) means √2
sqrt(30)
5*sqrt(2)
```
■

Definition

For vectors $\mathbf{x}=(x_1, x_2, \ldots, x_n)$, $\mathbf{y}=(y_1, y_2, \ldots, y_n)$ in $\mathbb{R}^n$,

$$x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

is called the **dot product** (or **Euclidean inner product**) of $\mathbf{x}$ and $\mathbf{y}$ and is denoted by $\mathbf{x} \cdot \mathbf{y}$. That is,

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

- Note that $\mathbf{x} \cdot \mathbf{x} = \| \mathbf{x} \|^2$

Example 2 Using the vectors $\mathbf{x}$ and $\mathbf{y}$ in **Example 1**, calculate $\mathbf{x} \cdot \mathbf{y}$.

Solution

$\mathbf{x} \cdot \mathbf{y} = 2 \cdot 3 + (-1) \cdot 2 + 3 \cdot 1 + 2 \cdot (-4) = 6 - 2 + 3 - 8 = -1$. □

Sage Copy the following code into http://sage.skku.edu to practice.

```
x=vector([2, -1, 3, 2])
y=vector([3, 2, 1, -4])
print x.inner_product(y)          # find the dot product
```

```
-1
```

■

Theorem 1.2.1

If $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ are vectors in $\mathbb{R}^n$ and k is a scalar, then we have the following:

(1) $\mathbf{x} \cdot \mathbf{x} \geq 0$,
(2) $\mathbf{x} \cdot \mathbf{x} = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$
(3) $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$
(4) $\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$
(5) $(k\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (k\mathbf{y}) = k(\mathbf{x} \cdot \mathbf{y})$

The proof of all the facts in above theorem are easy and users are encouraged to complete the same.

Theorem 1.2.2 [The Cauchy-Schwarz inequality]

For any two vectors $\mathbf{x}$, $\mathbf{y}$ in $\mathbb{R}^n$,

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \, \|\mathbf{y}\| .$$

Equality holds if and only if $\mathbf{x}$ and $\mathbf{y}$ are scalar multiples of one another (i.e. $\mathbf{x} = k\mathbf{y}$ for some scalar k).

- The Cauchy-Schwarz inequality is one of the most important inequalities in vector spaces. We will give a full details of this proof in section 9.2. This inequality implies $\frac{|\mathbf{x} \cdot \mathbf{y}|}{\| \mathbf{x} \| \ \| \mathbf{y} \|} \leq 1$ and $-1 \leq \frac{\mathbf{x} \cdot \mathbf{y}}{\| \mathbf{x} \| \ \| \mathbf{y} \|} \leq 1$ and which gives $\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\| \mathbf{x} \| \ \| \mathbf{y} \|}$ where $\cos\theta \in [-1, 1]$. This is a more generalized concept of the angle between two vectors, since these vectors can be matrices, polynomials, functions, etc.

Definition

For vectors $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ in $\mathbb{R}^n$

$$\mathbf{x} \cdot \mathbf{y} = \| \mathbf{x} \| \ \| \mathbf{y} \| \cos\theta, \ 0 \leq \theta \leq \pi,$$

where θ is called the **angle** between $\mathbf{x}$ and $\mathbf{y}$.

[Remark] **Parallelism and Orthogonality**

If $\mathbf{x} \cdot \mathbf{y} = 0$, then $\mathbf{x}$ is **orthogonal** to $\mathbf{y}$.
If $\mathbf{x}$ is a scalar multiple of $\mathbf{y}$ (i.e., $\mathbf{x} = k\mathbf{y}$ for some scalar k), then $\mathbf{x}$ is parallel to $\mathbf{y}$.

Definition

A vector $\mathbf{u}$ in R^n with a norm of 1, that is,

$$\| \mathbf{u} \| = 1$$

is called a **unit vector**. Additionally, if $\mathbf{x}$ and $\mathbf{y}$ are mutually orthogonal unit vectors, $\mathbf{x}$ and $\mathbf{y}$ are called **orthonormal vectors.**

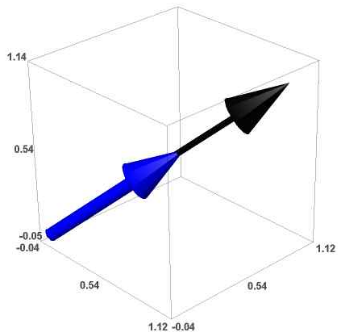


Figure 7

If $\mathbf{x}$ is a non zero vector, then $\mathbf{u} = \dfrac{\mathbf{x}}{\| \mathbf{x} \|}$, hence we have $\| \mathbf{u} \| = \left\| \dfrac{\mathbf{x}}{\| \mathbf{x} \|} \right\| = \dfrac{\| \mathbf{x} \|}{\| \mathbf{x} \|} = 1$.

Example 3

For two vectors $\mathbf{x} = (1,0,1,1)$ and $\mathbf{y} = (-1,0,0,1)$ in $\mathbb{R}^4$, establish orthogonality.

- http://matrix.skku.ac.kr/RPG_English/1-TF-inner-product.html

Solution

$\mathbf{x} \cdot \mathbf{y} = 1 \cdot (-1) + 0 \cdot 0 + 1 \cdot 0 + 1 \cdot 1 = 0$ □

Sage Copy the following code into http://sage.skku.edu to practice.

```
x=vector([1, 0, 1, 1])
y=vector([-1, 0, 0, 1])
print x.inner_product(y)
```

```
0    #orthogonal
```

■

Theorem 1.2.3 [Triangle Inequality for Vectors]

For any two vectors $\mathbf{x}$, $\mathbf{y}$ in $\mathbb{R}^n$. we have

$$\| \mathbf{x}+\mathbf{y} \| \le \| \mathbf{x} \| + \| \mathbf{y} \|$$

Equality holds if and only if $\mathbf{x}$ and $\mathbf{y}$ are *non-negative* scalar multiples of one another (i.e. $\mathbf{x} = k\mathbf{y}$ for some scalar $k \ge 0$).

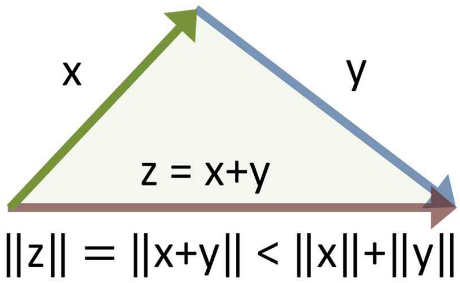


Figure 8

Geometrically, the sum of any length of any two sides of a triangle is greater than or equal to the third side. Look at the above figure.

Example 4 Using the vectors $\mathbf{x}$ and $\mathbf{y}$ from Example 1, verify that the triangle inequality holds.

Solution

$\mathbf{x}=(2,-1,3,2)$, $\mathbf{y}=(3,2,1,-4)$, $\| \mathbf{x} \| = \sqrt{4+1+9+4} = \sqrt{18} = 3\sqrt{2}$,
$\| \mathbf{y} \| = \sqrt{9+4+1+16} = \sqrt{30}$ and
$\mathbf{x}+\mathbf{y} = (2,-1,3,2)+(3,2,1,-4) = (5,1,4,-2)$. Hence
$\| \mathbf{x}+\mathbf{y} \| = \sqrt{25+1+16+4} = \sqrt{46}$.
So, $\|\mathbf{x}+\mathbf{y}\| = \sqrt{46} < \sqrt{18} + \sqrt{30} = \|\mathbf{x}\|+\|\mathbf{y}\|$. ■

Definition

For an arbitrary, non-zero vector $\mathbf{x}\,(\neq \mathbf{0}) \in \mathbb{R}^n$

$$\mathbf{u} = \frac{1}{\|\mathbf{x}\|}\mathbf{x}$$

is a unit vector. In $\mathbb{R}^n$, unit vectors of the form

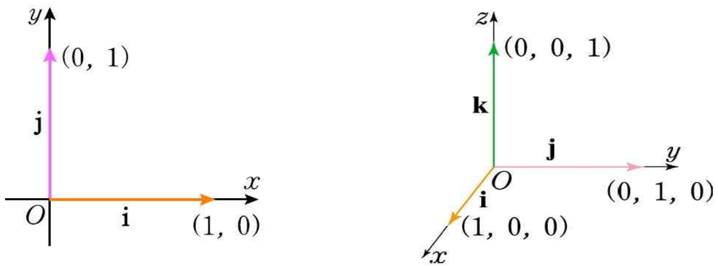
$$\mathbf{e}_1 = (1,0,0,\ldots,0),\ \mathbf{e}_2 = (0,1,0,\ldots,0),\ \ldots,\ \mathbf{e}_n = (0,0,0,\ldots,1)$$

are called **standard unit vectors** or coordinate vectors.

- If $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is an arbitrary vector in $\mathbb{R}^n$, using standard unit vectors, we can express $\mathbf{x}$ as follows:

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n.$$

- In $\mathbb{R}^2$ and $\mathbb{R}^3$, conventionally, the unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ along the rectangular coordinate axes are represented by $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$.



$$\mathbf{x} = x_1\mathbf{i} + x_2\mathbf{j},\quad \mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$$

$$\mathbf{x} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k},\quad \mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$$

Figure 9

<Figure 9 comes from Contemporary Linear algebra (3rd Edition) by Sang-Gu Lee, ISBN 978-89-6105-195-8, Kyungmoon Books(2009)>

1.3 Equations of Lines and Planes

- Reference video: http://youtu.be/4UGACWyWOgA http://youtu.be/YB976T1w0kE
- Practice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-1-Sec-1-3.html

In this section, we will derive vector equations of lines and planes in $\mathbb{R}^3$, and we will examine shortest distance problems related to these equations.

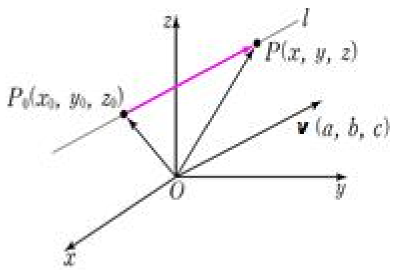
Point-Slope (Direction Vector) Equation of a Line

In $\mathbb{R}^3$, an equation of a line can be uniquely determined when a slope and a specified point on the line are given. If a line passes through the point $P_0(x_0, y_0, z_0)$ and is parallel to a vector $\mathbf{v} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$, then the vector $\overrightarrow{P_0P}$ is parallel to $\mathbf{v}$, where $P(x, y, z)$ is any point on the line.

That is, the line is a set of all points $P(x, y, z)$ that satisfies the following equation:

$$\overrightarrow{P_0P} = t\mathbf{v} \quad (t \in \mathbb{R})$$

Suppose $\overrightarrow{OP_0} = \mathbf{p}_0$ and $\overrightarrow{OP} = \mathbf{p}$. Then $\overrightarrow{P_0P} = \mathbf{p} - \mathbf{p}_0$. Hence we have $\mathbf{p} - \mathbf{p}_0 = t\mathbf{v}$.



That is, $\mathbf{p} = \mathbf{p}_0 + t\mathbf{v}$.

- **Vector equations**: $\mathbf{p} = \mathbf{p}_0 + t\mathbf{v}$, ($\mathbf{p} = \overrightarrow{OP}$, $\mathbf{p}_0 = \overrightarrow{OP_0}$)

- **Parametric equations**: In terms of coordinates, the above equations can be written as

$$x = x_0 + ta,\ \ y = y_0 + tb,\ \ z = z_0 + tc \qquad (-\infty < t < \infty).$$

- **Symmetric equations**: From the above parametric equations, it is easy to see that

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}\ (= t) \qquad (a,\ b,\ c \neq\ 0).$$

Example 1

Find vector, parametric and symmetric equations of the line that passes through the point $P(2, -1, 3)$ and is parallel to the vector $\mathrm{v} = (-3, 2, 4)$.

Solution

(1) The vector equation of the line is give by

$$x\mathrm{i} + y\mathrm{j} + z\mathrm{k} = 2\mathrm{i} - \mathrm{j} + 3\mathrm{k} + (-3\mathrm{i} + 2\mathrm{j} + 4\mathrm{k})t.$$

(2) The parametric equation is given by $\begin{cases} x = 2 - 3t \\ y = -1 + 2t \\ z = 3 + 4t \end{cases}$ $(-\infty < t < \infty)$.

(3) The symmetric equation is given by $\dfrac{x-2}{-3} = \dfrac{y+1}{2} = \dfrac{z-3}{4}$. ■

Example 2

Find parametric equations for the line that passes through the points $P(1,\ 1,\ -2)$ and $Q(4,\ -1,\ 0)$.

Solution Two points $P(1,\ 1,\ -2)$ and $Q(4,\ -1,\ 0)$ with position vectors r_0 and r_1 forms a vector

$$\overrightarrow{PQ} = \mathrm{v} = \mathrm{r}_1 - \mathrm{r}_0 = (4-1,\ -1-1,\ 0-(-2)) = (3,\ -2,\ 2)$$

and the vector equation $\mathrm{r} = \mathrm{r}_0 + t(\mathrm{r}_1 - \mathrm{r}_0)$ can be written as

$$\begin{aligned}(x,\ y,\ z) &= (1,\ 1,\ -2) + t(3,\ -2,\ 2) \\ &= (1+3t,\ 1-2t,\ -2+2t),\ (t \in \mathbb{R}).\end{aligned}$$

Thus, the parametric equations are:

$$x = 1 + 3t,\ y = 1 - 2t,\ z = -2 + 2t\ \ (-\infty < t < \infty) \qquad ■$$

Point-Normal Equation of Planes

A plane in $\mathbb{R}^3$ can be uniquely obtained by specifying a point $P_0(x_0, y_0, z_0)$ in the plane and a nonzero vector $\mathbf{n} = (a,b,c)$ that is perpendicular to the plane. The vector $\mathbf{n}$ is called the normal vector to the plane. If $P(x, y, z)$ is any point in this plane, then the $\overrightarrow{P_0P} = \mathbf{r} - \mathbf{r}_0$ is orthogonal to $\mathbf{n}$.
Hence by the property of the dot (inner) product.

$$\mathbf{n} \cdot \overrightarrow{P_0P} = (a, b, c) \cdot (x - x_0, y - y_0, z - z_0) = 0$$

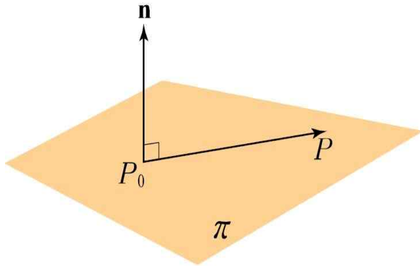


Figure 10

From this, we have

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

where a, b and c are not all zero.
This is called the point−normal equation of the plane through $P_0(x_0, y_0, z_0)$ with normal $\mathbf{n} = (a, b, c)$. The above equation can be simplified to $ax + by + cz = ax_0 + by_0 + cz_0 \ (= d)$.

Vector Equation of Planes

- Vector equations: A plane W in $\mathbb{R}^3$ can be uniquely obtained by passing through a point $\mathbf{x}_0 = P_0(x_0, y_0, z_0)$ and two nonzero vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in $\mathbb{R}^3$ that are not scalar multiples of one another.

  Let $\mathbf{x} = P(x, y, z)$ be any point on W. Then $\mathbf{x} - \mathbf{x}_0$ can be expressed as a linear combinations of $\mathbf{v}_1$ and $\mathbf{v}_2$. Look at the Figure 11.

$$\mathbf{x} - \mathbf{x}_0 = t_1\mathbf{v}_1 + t_2\mathbf{v}_2 \quad \text{or}$$
$$\mathbf{x} = \mathbf{x}_0 + t_1\mathbf{v}_1 + t_2\mathbf{v}_2 \quad (-\infty < t_1, t_2 < \infty)$$

where t_1 and t_2, called parameters, are in $\mathbb{R}$.

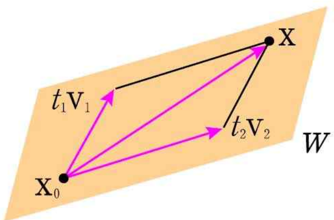


Figure 11

This is called a vector equation of the plane.

- Parametric equations: Let $\mathbf{x} = (x, y, z)$ be any point in the plane through $\mathbf{x}_0 = (x_0, y_0, z_0)$ that is parallel to the vectors $\mathbf{v}_1 = (a_1, b_1, c_1)$ and $\mathbf{v}_2 = (a_2, b_2, c_2)$. Then, we can express this in component form as

$$(x,\ y,\ z) = (x_0,\ y_0,\ z_0) + t_1(a_1,\ b_1,\ c_1) + t_2(a_2,\ b_2,\ c_2)$$

or

$$\begin{aligned} x &= x_0 + a_1t_1 + a_2t_2 \\ y &= y_0 + b_1t_1 + b_2t_2,\ \ (t_1,\ t_2 \in \mathbb{R}) \\ z &= z_0 + c_1t_1 + c_2t_2 \end{aligned}$$

These are called parametric equations of the plane.

Example 3

Find vector and parametric equations of the plane that passes through the three points: $P(4,-3,1)$, $Q(6,-4,7)$, and $R(1,2,2)$.

- http://matrix.skku.ac.kr/RPG_English/1-BN-11.html

Solution Let $\mathbf{x} = (x,\ y,\ z)$, $\mathbf{x}_0 = (4,\ -3,\ 1)$, $\mathbf{x}_1 = (6,\ -4,\ 7)$, and $\mathbf{x}_2 = (1,\ 2,\ 2)$. Then we have two vectors that parallel to the plane as

$$\mathbf{x}_1 - \mathbf{x}_0 = \overrightarrow{PQ} = (2,\ -1,\ 6),\ \ \mathbf{x}_2 - \mathbf{x}_0 = \overrightarrow{PR} = (-3,\ 5,\ 1).$$

Then, from our above definitions, we have

$$(x,\ y,\ z) = \mathbf{x}_0 + t_1(2,\ -1,\ 6) + t_2(-3,\ 5,\ 1),$$

which is a vector equation of the plane.

If we further simplify the above expression, we have

$$(x, y, z) = (4 + 2t_1 - 3t_2,\ -3 - t_1 + 5t_2,\ 1 + 6t_1 + t_2).$$

In particular, $x = 4 + 2t_1 - 3t_2$, $y = -3 - t_1 + 5t_2$, $z = 1 + 6t_1 + t_2$.

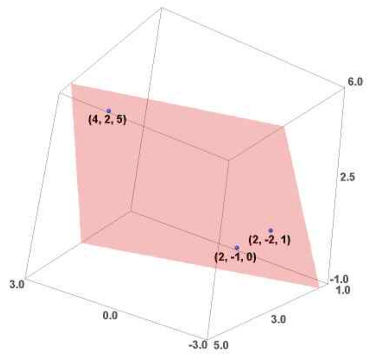
is the parametric equations of the plane. ■

[Remark] **Computer Simulation (A plane containing three points)**

- http://matrix.skku.ac.kr/2012-LAwithSage/interact/1/vec8.html

$$\det\begin{pmatrix} 1 & x & y & z \\ 1 & P_1 & Q_1 & R_1 \\ 1 & P_2 & Q_2 & R_2 \\ 1 & P_3 & Q_3 & R_3 \end{pmatrix} = 0$$

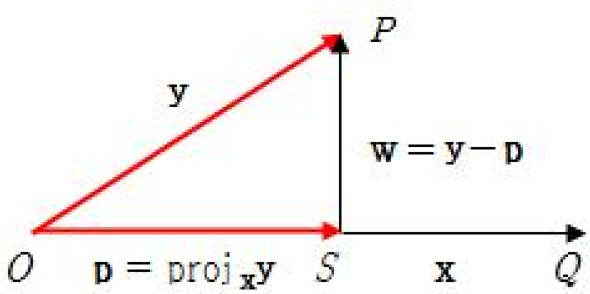
$$\Rightarrow -8x + 2y + 2z + 18 = 0$$



Vector Projection and Components

- Consider two vectors $\mathbf{x}$ and $\mathbf{y}$ with the same initial point O, represented by $\mathbf{x}=\overrightarrow{OQ}$ and $\mathbf{y}=\overrightarrow{OP}$. Let S be the foot of the perpendicular from P to the line containing $\overrightarrow{OQ}$. Then $\overrightarrow{OS}$ is called the **vector projection of y onto x** and is denoted by $\text{proj}_{\mathbf{x}}\mathbf{y}$.

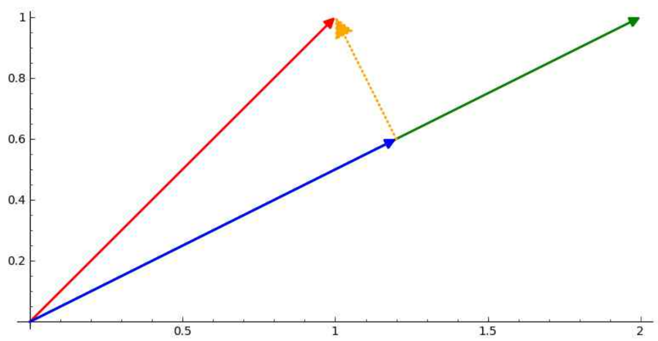
Here, the vector $\mathbf{w}=\overrightarrow{SP}$ is called the **component of y along x** (or the **scalar projection of y onto x**). Therefore, $\mathbf{y}$ can be written as $\mathbf{y}=\mathbf{p}+\mathbf{w}$.



Note that $\mathbf{p}$ is parallel to $\mathbf{x}$, hence $\mathbf{p}=t\,\mathbf{x}$ for some scalar t. Now $\mathbf{y}-\mathbf{p}$ is orthogonal to $\mathbf{x}$. Hence $\mathbf{x}\cdot(\mathbf{y}-\mathbf{p})=0$. This implies $t=(\mathbf{x}\cdot\mathbf{y})/(\mathbf{x}\cdot\mathbf{x})$. This gives the following results:

[Remark] Computer Simulation (Projection)

- http://matrix.skku.ac.kr/2012-LAwithSage/interact/1/vec3.html



Theorem 1.3.1 [Projection]

For vectors $\mathbf{x}\,(\neq \mathbf{0})$, $\mathbf{y}$ in $\mathbb{R}^3$, we have the following:

(1) $\operatorname{proj}_{\mathbf{x}}\mathbf{y} = t\,\mathbf{x} = \dfrac{(\mathbf{y}\cdot\mathbf{x})}{\mathbf{x}\cdot\mathbf{x}}\,\mathbf{x}$

(2) $D = \|\operatorname{proj}_{\mathbf{x}}\mathbf{y}\| = \dfrac{|\mathbf{y}\cdot\mathbf{x}|}{\|\mathbf{x}\|}$.

Example 4 For vectors $\mathbf{x}=(2,-1,3)$, $\mathbf{y}=(4,-1,2)$, find $\operatorname{proj}_{\mathbf{x}}\mathbf{y}$ (the vector projection of $\mathbf{y}$ onto $\mathbf{x}$) and the component of $\mathbf{y}$ along $\mathbf{x}$.

Solution

Since $\mathbf{y}\cdot\mathbf{x} = 15$, we have

$$\operatorname{proj}_{\mathbf{x}}\mathbf{y} = \frac{(\mathbf{y}\cdot\mathbf{x})}{\|\mathbf{x}\|^2}\mathbf{x} = \frac{15}{14}(2,-1,3) = \left(\frac{15}{7}, -\frac{15}{14}, \frac{45}{14}\right)$$

$$\mathbf{w} = \mathbf{y} - \operatorname{proj}_{\mathbf{x}}\mathbf{y} = (4,-1,2) - \left(\frac{15}{7}, -\frac{15}{14}, \frac{45}{14}\right) = \left(\frac{13}{7}, \frac{1}{14}, -\frac{17}{14}\right) \qquad \square$$

Sage Copy the following code into http://sage.skku.edu to practice.

```
a=vector([2, -1, 3])
b=vector([4, -1, 2])
ab=a.inner_product(b)
aa=a.inner_product(a)
p=ab/aa*a;w=b-p
print "p=", p
print "w=", w
```

```
p= (15/7, -15/14, 45/14)
w= (13/7, 1/14, -17/14)
```
■

Theorem 1.3.2 [Distance Between a Point and a Plane]

For a point $P_0(x_0, y_0, z_0)$ and a plane $\pi : ax+by+cz+d=0$, the distance D from the point to the plane is given by

$$D = \frac{|ax_0+by_0+cz_0+d|}{\sqrt{a^2+b^2+c^2}}.$$

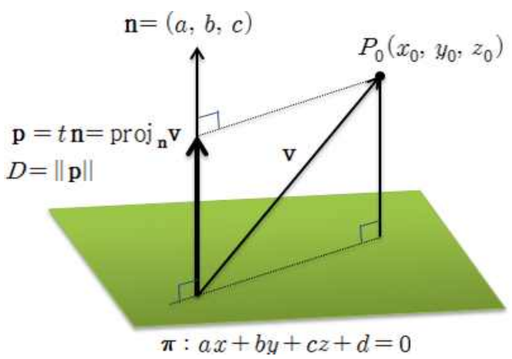


Figure 12

Note that the distance of the point P_0 from the orthogonal projection of the vector $\mathbf{v} = (x_0, y_0, z_0)$ onto the plane $ax+by+cz+d=0$. This distance is same as the orthogonal projection of the vector (x_0, y_0, z_0) onto the normal vector $\mathbf{n} = (a, b, c)$ to the plane. See the Figure 12. It is as easy exercise to verify that the orthogonal projection of $\mathbf{v}$ onto $\mathbf{n}$ is given by the formula D above.

Example 5 Find the distance D from the point $P(3,-1,2)$ to the plane $x+3y-2z-6=0$.

- http://matrix.skku.ac.kr/RPG_English/1-B1-point-plane-distance.html

Solution

$$\mathbf{p} = \text{proj}_{\mathbf{n}}\mathbf{v} = t\mathbf{n} = \frac{\mathbf{v}\cdot\mathbf{n}}{\mathbf{n}\cdot\mathbf{n}}\mathbf{n}$$

Here, $\mathbf{n}=(1,\ 3,\ -2)$, $\mathbf{v}=(3,\ -1,\ 2)$, and $d=-6$, so

$$D=\|\mathbf{p}\| = \|\text{proj}_{\mathbf{n}}\mathbf{v}\| = \frac{|3+3(-1)-2(2)-6|}{\sqrt{1^2+3^2+(-2)^2}} = \frac{10}{\sqrt{14}} = \frac{2}{7}\sqrt{14}$$ □

Sage Copy the following code into http://sage.skku.edu to practice.

```
n=vector([1, 3, -2])
v=vector([3, -1, 2]);d=-6
vn=v.inner_product(n)
nn=n.norm()
Dist=abs(vn+d)/nn
print Dist
```

5/7*sqrt(14) # $\frac{10}{\sqrt{14}}=\frac{2}{7}\sqrt{14}$ ■

[2014 Seoul International Congress of Mathematicians] http://www.icm2014.org/kr

Chapter 1 Exercises

- http://matrix.skku.ac.kr/LA-Lab/index.htm
- http://matrix.skku.ac.kr/knou-knowls/cla-sage-reference.htm

Problem 1 For points $P_1 = (5, -2, 1)$, $P_2 = (2, 4, 2)$, find the vector $\overrightarrow{P_1P_2}$.

Problem 2 What is the initial point of the vector $\mathbf{x} = (1, 1, 3)$ with terminal point $B(-1, -1, 2)$?

Problem 3 For vectors $\mathbf{u} = (-3, 1, 2, 4, 4)$, $\mathbf{v} = (4, 0, -8, 1, 2)$, and $\mathbf{w} = (6, -1, -4, 3, -5)$, compute the following:

$$(2\mathbf{u} - 7\mathbf{w}) - (8\mathbf{v} + \mathbf{u})$$

Problem 4 Using the same $\mathbf{u}, \mathbf{v}, \mathbf{w}$ from above, find the vector $\mathbf{x}$ that satisfies the following:

$$2\mathbf{u} - \mathbf{v} + \mathbf{x} = 7\mathbf{x} + \mathbf{w}$$

Problem 5 For vectors $\mathbf{x} = (-1, -2, 3)$, $\mathbf{y} = (3, -2, -1)$ calculate $\cos\theta$, where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$.

Problem 6 Find the distance between the two points $P(-1, 2, 1)$ and $Q(-3, -4, 5)$.

Problem 7 For vectors $\mathbf{x}=(a,\ 2,\ -1,\ a)$, $\mathbf{y}=(-a,\ -1,\ 3,\ 6)$, find the real number that such that $\mathbf{x}\cdot\ \mathbf{y}=0$.

Problem 8 Find a vector equation of the line between the two points $P(-5,\ 1,\ 3)$, and $Q(2,\ -3,\ 4)$.

Problem 9 Find a normal vector perpendicular to the plane $z=-7x+y+4$.

Problem 10 [Projection] For $\mathbf{x}=(2,\ -1,\ 3)$ and $\mathbf{y}=(4,\ -1,\ 2)$, find the scalar projection and vector projection of $\mathbf{y}$ onto $\mathbf{x}$.

Solution $\mathrm{proj}_{\mathbf{x}}\mathbf{y}=\dfrac{\mathbf{y}\cdot\ \mathbf{x}}{\mathbf{x}\cdot\ \mathbf{x}}\mathbf{x}=\dfrac{15}{14}(2,-1,3)=(\dfrac{15}{7},-\dfrac{15}{14},\dfrac{45}{14})$

$\mathbf{w}=\mathbf{y}-\mathrm{proj}_{\mathbf{x}}\mathbf{y}=(4,-1,2)-(\dfrac{15}{7},-\dfrac{15}{14},\dfrac{45}{14})=(\dfrac{13}{7},\dfrac{1}{14},-\dfrac{17}{14})$

Sage :

```
a=vector([2, -1, 3])
b=vector([4, -1, 2])
ab=a.inner_product(b)
aa=a.inner_product(a)
p=ab/aa*a;w=b-p
print "p=", p
print "w=", w
```

```
p= (15/7, -15/14, 45/14)
w= (13/7, 1/14, -17/14)
```

■

Problem P1 [Discussion] Vectors with the same magnitude and direction are considered to be equivalent. However, in a vector space, discuss the relationship between vectors with the same slope but expressed with different equations.

Problem P2 [Discussion] For vectors $\mathbf{v}_1=\left(\frac{2}{3},\ \frac{1}{3},\ \frac{2}{3}\right)$ and $\mathbf{v}_2=\left(\frac{1}{3},\ \frac{2}{3},\ -\frac{2}{3}\right)$, check if $\mathbf{v}_1$ and $\mathbf{v}_1$ are orthonormal vectors, and find a third vector $\mathbf{v}_3$ such that $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ are all orthonormal to one another.

Solution $\|v_1\| = \sqrt{(\frac{2}{3})^2 + (\frac{1}{3})^2 + (\frac{2}{3})^2} = 1$, $\|v_2\| = \sqrt{(\frac{1}{3})^2 + (\frac{2}{3})^2 + (-\frac{2}{3})^2} = 1$

$v_1 \cdot v_2 = \frac{2}{9} + \frac{2}{9} - \frac{4}{9} = 0$ => v_1 and v_2 are orthonormal.

Let $v_3 = (a, b, c)$ such that $\|v_3\| = \sqrt{a^2 + b^2 + c^2} = 1$, $v_1 \cdot v_3 = \frac{2}{3}a + \frac{1}{3}b + \frac{2}{3}c = 0$,

$v_2 \cdot v_3 = \frac{1}{3}a + \frac{2}{3}b - \frac{2}{3}c = 0$. => $a = \frac{2}{3}$, $b = -\frac{2}{3}$, $c = -\frac{1}{3}$

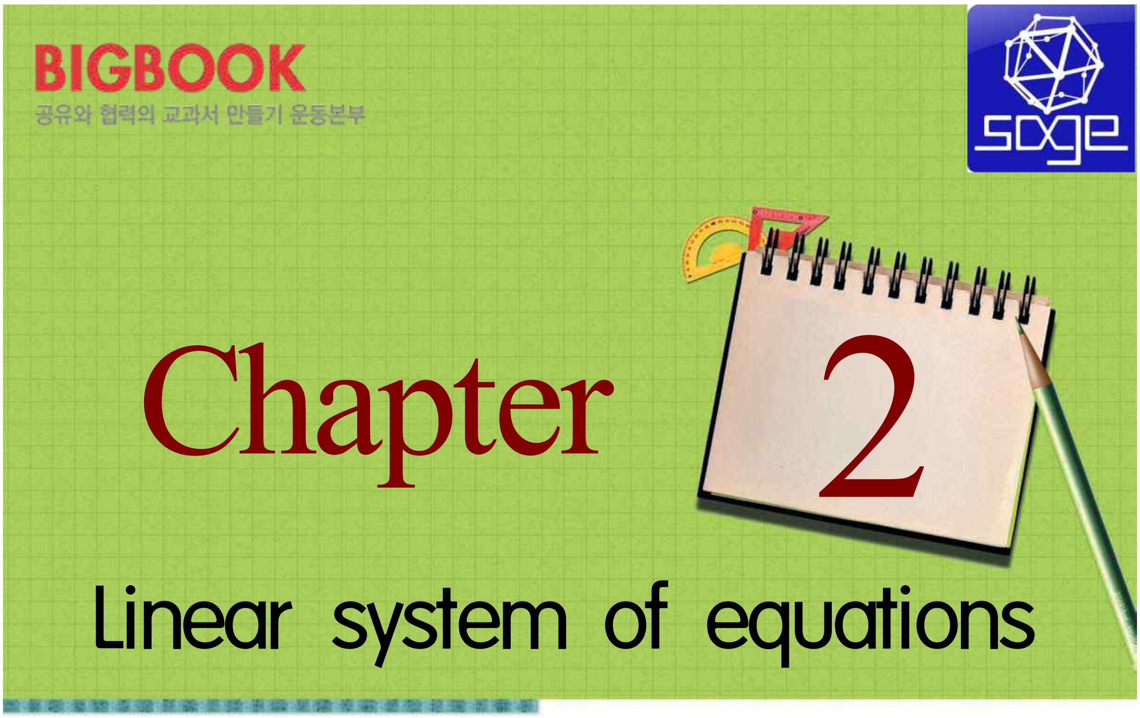
This shows $v_3 = \left(\frac{2}{3}, -\frac{2}{3}, -\frac{1}{3}\right)$ ■

[Digital Library of Math Textbooks in 60's at SKKU]

http://matrix.skku.ac.kr/2012-e-Books/index.htm

<1884~1910 Math books written by Korean authors>

http://www.hpm2012.org/Proceeding/Exhibition/E2.pdf



A system of linear equations and its solution is one of the most important problems in Linear Algebra. A linear system with thousands of variables occurs in natural and social sciences, engineering, as well as traffic problems, weather forecasting, decision-making, etc. Even differential equations concerning derivatives such as velocity and acceleration can be solved by transforming them into a linear system.

In Linear Algebra, a solution of a linear system is obtained by Gauss elimination method or with determinants. In Chapter 2, we consider a geometric meaning of the solution of a linear system and its solution, and investigate some applications of a linear system.

2.1 Linear system of equations

- Reference video: http://youtu.be/CiLn1F2pmvY, http://youtu.be/AAUQvdjQ-qk
- Practice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-2-Sec-2-1.html

The theory of linear systems is the basis and a fundamental part of linear algebra, a subject which is used in most parts of modern mathematics. Computational algorithms for finding the solutions are an important part of numerical linear algebra, and play a prominent role in engineering, physics, chemistry, computer science, and economics. In this section, we study the process of finding solutions of linear system of equations and its geometric meanings.

Definition [Linear equations]

Let b and $a_1, a_2, \dots, a_n$ be real numbers. A linear equation with unknowns $x_1, \dots, x_n$ is of the following form:

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

In other words, a linear equation consists of variables of degree 1 and a constant.

Example 1

Equations $2x_1 - 3x_2 + 1 = x_1$, $x_2 = 2(\sqrt{5} - x_1) - x_3$ can be written as $x_1 - 3x_2 = -1$, $2x_1 + x_2 + x_3 = 2\sqrt{5}$ and they are linear. But $2x_1 - 3x_2 = x_1x_2$, $x_2 = 3\sqrt{x_1} - 1$, $x_1 + \sin x_2 = 0$ are not linear.

Definition [Linear system of equations]

In general, a set of m linear equations with unknowns $x_1, x_2, \dots, x_n$

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

is called a system of linear equations. If constants $b_1, b_2, \dots, b_m$ are all zeroes, it is called a homogeneous system of linear equations.

Definition [Solutions of a linear system]

Suppose that unknowns $x_1, x_2, \dots, x_n$ in a linear system are substituted by $s_1, s_2, \dots, s_n$ respectively and each equation is satisfied. Then $s_1, s_2, \dots, s_n$ is called a solution of a linear system. For example, given a linear system

$$\begin{aligned} 4x_1 - x_2 + 3x_3 &= -1 \\ 3x_1 + x_2 + 9x_3 &= -4 \end{aligned} \qquad (2)$$

One can substitute x_1, x_2, x_3 as $1, 2, -1$, respectively, and it satisfies equation (2). Hence $(1, 2, -1)$ is a solution. In general, if there is a solution of a linear system, it is called **consistent** and is called **inconsistent** otherwise.

- The set of all solutions of a linear system is called a solution set. Two linear systems with the same solution set are called **equivalent.**

[Remark] **Solution** (linear system with two unknowns)

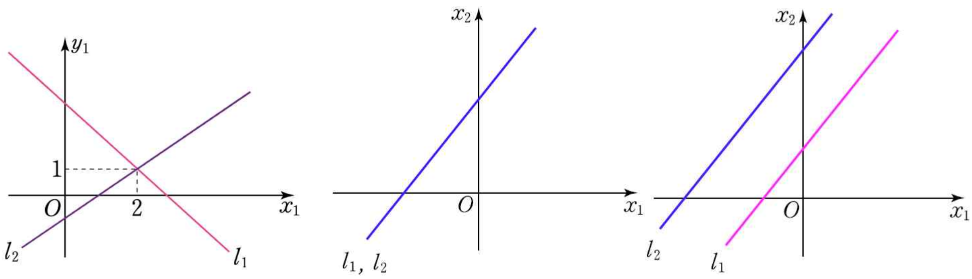
In general, a given linear system satisfies one and only of the following.

(1) a unique solution

(2) infinitely many solutions

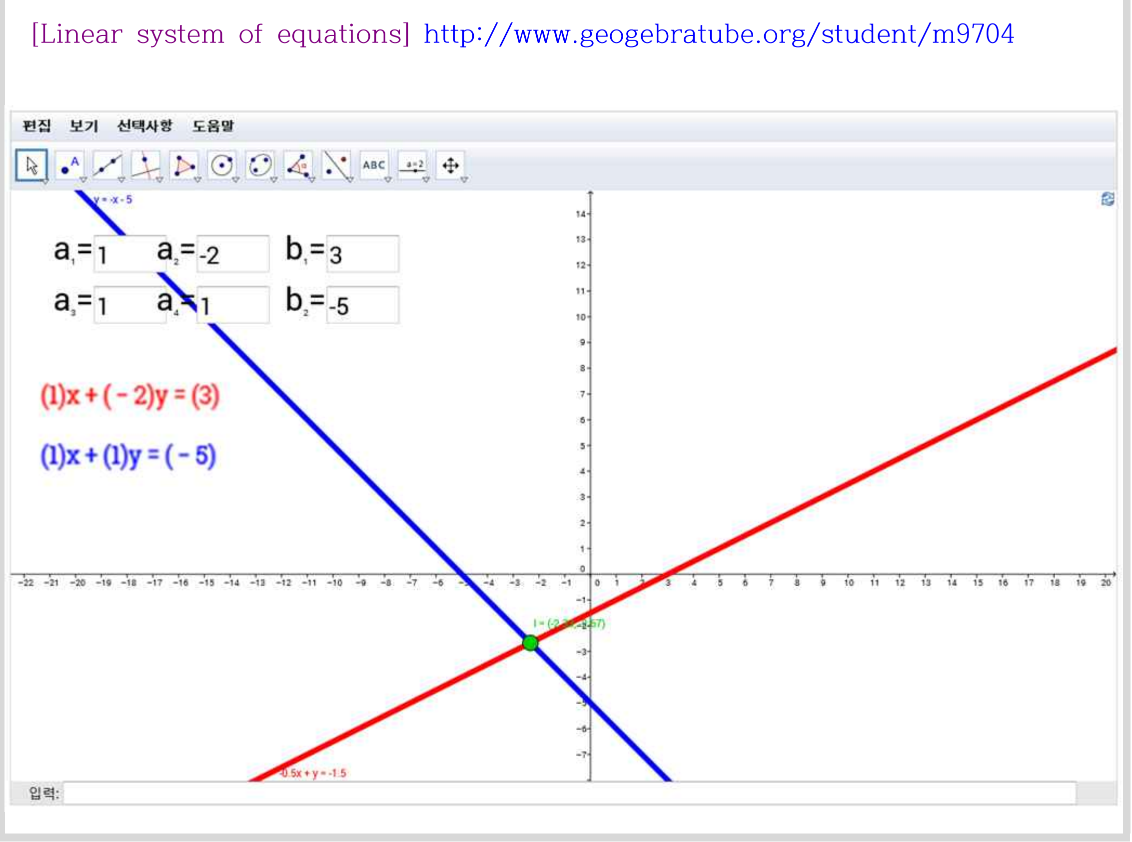
(3) no solution

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

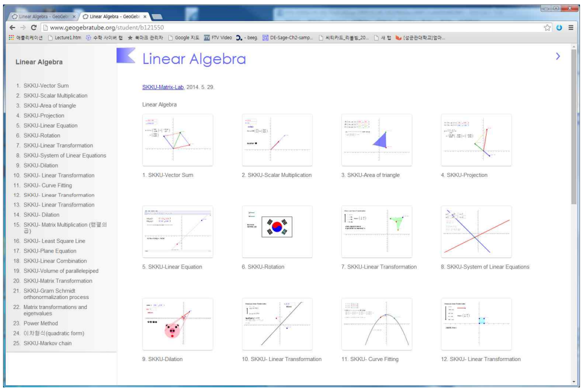


[Remark] Computer simulation

[Linear system of equations] http://www.geogebratube.org/student/m9704



[Remark] Linear algebra with Geogebra:



http://www.geogebratube.org/student/b121550

Remark: (i) If there is one linear equation in three variables then it has infinitely many solutions. (ii) If there are two linear equations in three variables then, it either it has no solution (when the two planes are parallel) or it has infinitely many solutions which is the points of line of intersection of the two planes. (iii) In case of three linear equations in three variables, all possibilities can occur.

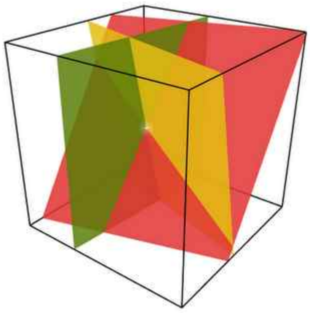
Example 2 Describe all possible solution sets in R^3 of a linear system with three equations and three unknowns.

Solution

One can show that there are three possibilities by a geometric method. Let us denote each equation by a plane H_1, H_2, H_3 respectively.

① It has a unique solution.

Three planes meets in a unique point.



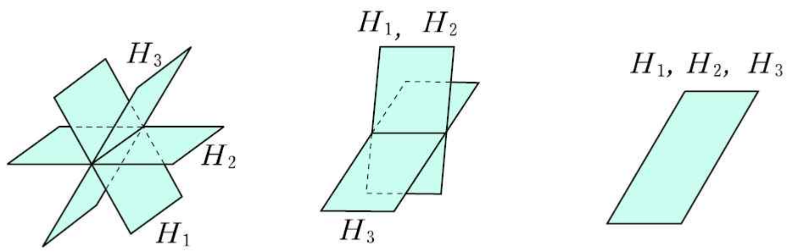
[Ex] $$\begin{aligned} x+y+2z&=-3\\ x+2y+3z&=-4\\ x+y+3z&=-6 \end{aligned}$$

② It has infinitely many solutions.

[Ex] (1) $$\begin{aligned} x+y+z&=1\\ 2x+2y+z&=3\\ x+2y+2z&=4 \end{aligned}$$

(2) $$\begin{aligned} x+y+z&=1\\ 2x+2y+2z&=2\\ 2x+y+3z&=4 \end{aligned}$$

(3) $$\begin{aligned} x+y+z&=1\\ 2x+2y+2z&=2\\ 3x+3y+3z&=3 \end{aligned}$$



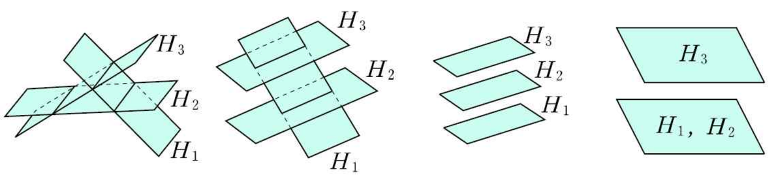
③ It has no solution. (It is called 'inconsistent').

[Ex] (1) $$\begin{aligned} x+y+z&=1\\ x\quad\;\; +z&=1\\ y\quad\;\;&=3 \end{aligned}$$

(2) $$\begin{aligned} x+y+z&=1\\ x+y+z&=2\\ 2x+y+z&=3 \end{aligned}$$

(3) $$\begin{aligned} x+y+z&=1\\ x+y+z&=2\\ x+y+z&=3 \end{aligned}$$

(4) $$\begin{aligned} x+y+z&=1\\ 2x+2y+2z&=2\\ 2x+2y+2z&=3 \end{aligned}$$



Example 3 Solve the following linear system.

$$\begin{aligned} x_1 - 2x_2 + x_3 + 3x_4 - x_5 &= 0 \\ x_3 - 5x_4 + 2x_5 &= 1 \\ x_4 + \ \ x_5 &= 2 \end{aligned}$$

Solution Since there are five unknowns and three equations, assign to the any two unknowns arbitrary real numbers. Rearranging each equation, we get

$$\begin{aligned} x_1 + x_3 + 3x_4 &= 2x_2 + x_5 \\ x_3 - 5x_4 &= 1 - 2x_5 \\ x_4 &= 2 - x_5 \end{aligned}$$

Substitute $x_2 = s$, $x_5 = t$ (s, t are arbitrary real numbers) to get

$$\begin{aligned} x_4 &= 2 - x_5 = 2 - t \\ x_3 &= 1 - 2x_5 + 5x_4 = 1 - 2t + 5(2-t) = 11 - 7t \\ x_1 &= 2x_2 + x_5 - x_3 - 3x_4 = 2s + t - (11 - 7t) - 3(2 - t) \\ &= -17 + 2s + 11t \end{aligned}$$

Therefore, the solution of a given linear system is

$$x_1 = -17 + 2s + 11t, \quad x_2 = s, \quad x_3 = 11 - 7t, \quad x_4 = 2 - t, \quad x_5 = t$$

(s, t are arbitrary real numbers).

The solution set is

$$S = \{(x_1, x_2, x_3, x_4, x_5) = (-17 + 2s + 11t,\ s, 11 - 7t, 2 - t,\ t) \mid s, t \in R\}.$$

Thus this system has infinitely many solutions. ■

Definition [Matrix]

An array (or rectangle) consisting of real (or complex) numbers is called a matrix, and each number is called an entry.

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

The row $[a_{i1} \quad a_{i2} \quad \cdots \quad \cdots \quad a_{in}]\,(1 \leq i \leq m)$ of matrix A is called the i**-th row of** A, and the column

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \quad (1 \leq j \leq n)$$

of A is called the j**-th column of** A. A matrix with m rows and n columns is called a **size** $m \times n$ matrix, and if $m = n$, it is called a square matrix of order n.

- Let $A_{(i)}$ **denote the** i**th row of** A**, and** $A^{(j)}$ **denote the** j**the column of** A. Therefore we can write A as follows.

$$A = \begin{bmatrix} A_{(1)} \\ \vdots \\ A_{(m)} \end{bmatrix} = [A^{(1)} \,|\, \cdots \,|\, A^{(n)}]$$

The entry a_{ij} of a matrix A is also called the (i, j) entry of A, and the entries $a_{11}, a_{22}, \ldots, a_{nn}$ of a matrix of order n are called main diagonal entries. Matrix (2) can be written as the (i, j) entries as follows.

$$A = [a_{ij}]_{m \times n} \text{ or } A = [a_{ij}]$$

Example 4 Consider matrices

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

$$C = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 1 & -3 \\ 2 & 0 & 1 \\ 5 & -1 & 2 \end{bmatrix}, \quad E = [2], \quad F = [-1 \quad 0 \quad 3].$$

A is a 2×3 matrix, and $a_{13}=-1$, $a_{22}=3$. B is a 2×2 matrix, and $b_{11}=1$, $b_{22}=-3$, and C, D, E, F are 3×1, 3×3, 1×1, 1×3 matrix respectively. The main diagonal entries of D are $d_{11}=1$, $d_{22}=0$, $d_{33}=2$, and E is also written as $E=[2]=2$. ■

Definition **[Coefficient matrix and augmented matrix of a linear system]**

For a linear system with n unknowns and m linear equations

$$\begin{aligned} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n &= b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n &= b_m, \end{aligned} \qquad (3)$$

let

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{b}=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

then Equation (3) can be written as

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

The matrix A is called the coefficient matrix of Equation (3) and the matrix obtained from A and $\mathbf{b}$

$$[A \,\vdots\, \mathbf{b}]=\left[\begin{array}{cccc:c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

is called the augmented matrix of Equation (3).

Example 5

Find the augmented matrix of the following linear system of equations.

$$\begin{aligned} x+\ \ y+2z&=9\\ 2x+4y-3z&=1\\ 3x+6y-5z&=0 \end{aligned}$$

Solution Let A be the coefficient matrix, $\mathbf{x}$ the unknown, and $\mathbf{b}$ the constant, then

$$A=\begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & -3 \\ 3 & 6 & -5 \end{bmatrix},\ \mathbf{x}=\begin{bmatrix} x \\ y \\ z \end{bmatrix},\quad \mathbf{b}=\begin{bmatrix} 9 \\ 1 \\ 0 \end{bmatrix}$$

Hence we have

$$A\mathbf{x}=\mathbf{b} \quad \Leftrightarrow \quad \begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & -3 \\ 3 & 6 & -5 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} 9 \\ 1 \\ 0 \end{bmatrix}$$

Its augmented matrix is

$$[A \,\vdots\, \mathbf{b}]=\left[\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{array}\right]$$

□

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(3, 3, [1,1,2,2,4,-3,3,6,-5])          # 3x3 matrix
b=vector([9,1,0])                               # constant vector
print A.augment(b,subdivide=True)               # augmented matrix
```

```
[ 1  1  2 |  9 ]
[ 2  4 -3 |  1 ]
[ 3  6 -5 |  0 ]
```

■

2.2 Gaussian elimination and Gauss-Jordan elimination

- Reference video: http://youtu.be/jnC66zvqHJI, http://youtu.be/HSm69YigRr4
- Practice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-2-Sec-2-2.html

Gaussian elimination (also known as row reduction) is an algorithm for solving systems of linear equations. It is usually understood as a sequence of operations performed on the associated matrix of coefficients. Using row operations to convert a matrix into reduced row echelon form is sometimes called Gauss-Jordan elimination. Linear system of equations can be easily solved by using Gauss-Jordan elimination.

- Solving a linear system: using elimination method:

$$\begin{cases} 2x+3y=1 \\ x-2y=4 \end{cases} \Rightarrow \text{Multiplying 2 on the second equation, } \begin{cases} 2x+3y=1 \\ 2x-4y=8 \end{cases}$$

$\Rightarrow$ Subtracting the second equation from the first equation, $\begin{cases} 2x+3y= & 1 \\ 7y= & -7 \end{cases}$

$\Rightarrow$ Dividing the second equation by 7, $\begin{cases} 2x+3y= & 1 \\ y= & -1 \end{cases}$

$\Rightarrow$ Substituting $y=-1$ in the first equation, $\begin{cases} 2x= & 4 \\ y= & -1 \end{cases} \Rightarrow \begin{cases} x= & 2 \\ y= & -1 \end{cases}$

- The following **operations** do not change the solution set.

(1) Exchange two equations. $R_i \leftrightarrow R_j$

(2) Multiply a row by a nonzero real number. kR_i

(3) Add a nonzero multiple of a row to another row. $kR_i + R_j \rightarrow R_j$

These are called **Elementary Row Operations (ERO)**.

Example 1

In the following procedure, the left side shows solving a linear system directly, and the right side shows solving it using an augmented matrix.

$$\begin{aligned} x + y+2z &= 9 \\ 2x+4y-3z &= 1 \\ 3x+6y-5z &= 0 \end{aligned} \qquad \left[\begin{array}{ccc:c} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{array}\right]$$

Add the multiplication of the first equation by -2 to the second equation.

$$\begin{aligned} x + \quad y + 2z &= \quad 9 \\ 2y - 7z &= -17 \\ 3x + 6y - 5z &= \quad 0 \end{aligned} \quad \xrightarrow{(-2)R_1 + R_2} \quad \left[\begin{array}{ccc:c} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 3 & 6 & -5 & 0 \end{array}\right]$$

Add the multiplication of the first equation by -3 to the third equation.

$$\begin{aligned} x + \quad y + \ 2z &= \quad 9 \\ 2y - \ 7z &= -17 \\ 3y - 11z &= -27 \end{aligned} \quad \xrightarrow{(-3)R_1 + R_3} \quad \left[\begin{array}{ccc:c} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 0 & 3 & -11 & -27 \end{array}\right]$$

Multiply the second equation by 1/2 to get

$$\begin{aligned} x + y + 2z &= 9 \\ y - \frac{7}{2}z &= -\frac{17}{2} \\ 3y - 11z &= -27 \end{aligned} \quad \xrightarrow{\frac{1}{2}R_2} \quad \left[\begin{array}{ccc:c} 1 & 1 & 2 & 9 \\ 0 & 1 & \frac{7}{2} & -\frac{17}{2} \\ 0 & 3 & -11 & -27 \end{array}\right]$$

Add the multiplication of the second equation by -3 to the third equation.

$$\begin{aligned} x + y \ + 2z &= 9 \\ y - \frac{7}{2}z &= -\frac{17}{2} \\ -\frac{1}{2}z &= -\frac{3}{2} \end{aligned} \quad \xrightarrow{(-3)R_2 + R_3} \quad \left[\begin{array}{ccc:c} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & -\frac{1}{2} & -\frac{3}{2} \end{array}\right]$$

Multiply the third equation by -2.

$$\begin{aligned} x + y + 2z &= \quad 9 \\ y - \frac{7}{2}z &= -\frac{17}{2} \\ z &= \quad 3 \end{aligned} \quad \xrightarrow{(-2)R_3} \quad \left[\begin{array}{ccc:c} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Thus the system reduces to
$$\begin{aligned} z &= 3 \\ y &= \frac{7}{2}z - \frac{17}{2} = 2 \\ x &= 9 - y - 2z = 1 \end{aligned}$$

Now substituting $z = 3$ in the second equation, we get $y = 2$. Substituting $y = 2$ and $z = 3$ in the first equation, we get $x = 1$. Hence the solution is $x = 1,\ y = 2,\ z = 3$. ■

Definition **[Row echelon form(REF) and reduced row echelon form(RREF)]**

When an $m \times n$ matrix E satisfies the following 3 properties, it is called a row echelon form (**REF**).

(1) If there is a row consisting of only 0's, it is placed on the bottom position.

(2) The first nonzero entry appearing in each row is 1. This 1 is called a leading entry.

(3) If there is a leading entry in both the ith row and the $(i+1)$ row, the leading entry in the $(i+1)$th row is placed on the right of the leading entry in the ith row.

If matrix E is a REF and satisfies the following property, E is called a reduced row echelon form (**RREF**).

(4) If a column contains the leading entry of some row, then all the other entries of that column are 0

Example 2 The following are all REF.

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

Example 3 Consider matrices

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

Since matrices A, B, C do not satisfy the above properties (1), (2), (3) respectively, they are not REF.

Example 4 The following are all RREF.

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

[Remark]

Below are a general form of a REF and its corresponding RREF(here * is any number).

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

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

Definition **[Elementary Row Operation(ERO)]**

Given an $m \times n$ matrix A, the following operations are called elementary row operation (ERO).

E1: Exchange the ith row and the jth row of A. $R_i \leftrightarrow R_j$

E2: Multiply the ith row of A by a nonzero constant k. kR_i

E3: Add the multiplication of the ith row of A by k to the jth row. $kR_i + R_j$

- EROs transform a given matrix into REF and RREF.

Definition [Row Equivalent]

If B is obtained from a matrix A by elementary row operations, A and B are row equivalent.

Example 5 The following are equivalent.

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

Finding REF and RREF

- For $A=\begin{bmatrix}0&0&-2&0&7&12\\2&4&-10&6&12&28\\2&4&-5&6&-5&-1\end{bmatrix}$, find REF and RREF by applying ERO's.

Find a column whose entries are not all zero and which is located in left-most position.

$$\begin{bmatrix}0&0&-2&0&7&12\\2&4&-10&6&12&28\\2&4&-5&6&-5&-1\end{bmatrix}$$

(In this case, it is the first column)

Swap the first row with some other row below to guarantee that a_{11} is not zero.

$$\begin{bmatrix}2&4&-10&6&12&28\\0&0&-2&0&7&12\\2&4&-5&6&-5&-1\end{bmatrix}$$

Swap 1st and 2nd row

(In this case, $a_{21}=2$ became $a_{11}{}'$. This $a_{21}=2$ is call a *pivot*.)

Divide the 1st row by 2 to make the pivot entry $=1$.

$$\begin{bmatrix}1&2&-5&3&6&14\\0&0&-2&0&7&12\\2&4&-5&6&-5&-1\end{bmatrix}$$

Multiply $\dfrac{1}{2}$ to the 1st row.

Eliminate all other entries in the 1st column by subtracting suitable multiples of the 1st row from the other rows.
(Use elementary row operations).

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}$$

Eliminate $a_{13} = 2$ in the 1st column by subtracting -2 multiple of the 1st row from the 3rd row.

Continue steps 1, 2, 3, 4 for the remaining rows except the 1st row .

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}$$

Find a column whose entries are not all zero and which is located in the left-most position (excluding the 1st column).

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}$$

Since the leading entry is not 1, follow step 3.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix}$$

Eliminate $a_{33}'' = 5$ in the 3rd column by subtracting -5 multiple of the 2nd row from the 3rd row.

Continue steps 1, 2, 3 for the rows except the 1st and 2nd row .

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix}$$

Find a column whose entries are not all zero (excluding the 1st and 2nd rows).

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

Since there is a row whose entries are not all zero, follow step 3.

Therefore we have REF of A as follows.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

Furthermore, we get the RREF of A from the above REF by making nonzero a_{ii}"to be 1 by suitable multiples of each row.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

Add the $\frac{7}{2}$ multiple of 3rd row to 2nd row.

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

Add the -6 multiple of 3rd row to 1st row.

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 & 7 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

Add the 5 multiple of 3rd row to 1st row.

Now we have the RREF of A.

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 & 7 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \qquad \square$$

http://www.math.odu.edu/~bogacki/cgi-bin/lat.cgi?c=rref

Example 6 Find the RREF of A.

$$A=\begin{bmatrix} 1 & 1 & 1 & 4 & 4 \\ 2 & 3 & 4 & 9 & 16 \\ -2 & 0 & 3 & -7 & 11 \end{bmatrix}$$

Solution http://matrix.skku.ac.kr/RPG_English/2-MA-RREF.html
http://matrix.skku.ac.kr/2014-Album/MC.html

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(3, 5, [1,1,1,4,4,2,3,4,9,16,-2,0,3,-7,11])     # 3x5 matrix input
print A.rref()                                          # A's RREF
```

```
[ 1  0  0  2 -1]
[ 0  1  0  3  2]
[ 0  0  1 -1  3]
```
■

Theorem 2.2.1

Two linear systems whose augmented matrices are row equivalent are equivalent (that is, they have the same solution sets.)

- **Gauss elimination**: This is a method to transform the augmented matrix of a linear system into REF.

Example 7

Solve the following by the Gauss elimination.

$$\begin{aligned} x+2y+3z &= 9 \\ 2x-y+z &= 8 \\ 3x \quad\quad -z &= 3 \end{aligned}$$

Solution Its augmented matrix is $\begin{bmatrix} 1 & 2 & 3 & \vdots & 9 \\ 2 & -1 & 1 & \vdots & 8 \\ 3 & 0 & -1 & \vdots & 3 \end{bmatrix}$ and its REF by EROs is $\begin{bmatrix} 1 & 2 & 3 & : & 9 \\ 0 & 1 & 1 & : & 2 \\ 0 & 0 & 1 & : & 3 \end{bmatrix}$. Therefore, since the corresponding linear system of the above augmented matrix is

$$\begin{cases} x+2y+3z=9 \\ \quad\quad y+z=2 \\ \quad\quad\quad z=3 \end{cases} \text{ ,i.e, } \begin{cases} x=3 \\ y=2-z=-1 \\ z=3 \end{cases}$$

The solution is $x=2,\ y=-1,\ z=3$. ■

- **Gauss-Jordan elimination**: This is a method to transform the augmented matrix of a linear system into RREF.

Example 8 Solve the following system using the Gauss-Jordan elimination.

$$\begin{aligned} x_1+3x_2-2x_3 \quad\quad\quad +2x_5 \quad\quad &= 0 \\ 2x_1+6x_2-5x_3-2x_4+4x_5-3x_6 &= -1 \\ 5x_3+10x_4 \quad\quad +15x_6 &= 5 \\ 2x_1+6x_2 \quad\quad +8x_4+4x_5+18x_6 &= 6 \end{aligned}$$

Solution We will use Sage to solve this. □

- http://matrix.skku.ac.kr/RPG_English/2-VT-Gauss-Jordan.html

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix([[1,3,-2,0,2,0],[2,6,-5,-2,4,-3],[0,0,5,10,0,15], [2,6,0,8,4,18]])
b=vector([0,-1,5,6])
print A.augment(b).rref()
```

```
[  1   3   0   4   2   0   0]
[  0   0   1   2   0   0   0]
[  0   0   0   0   0   1 1/3]
[  0   0   0   0   0   0   0]
```

Its corresponding linear system is

$$\begin{cases} x_1+3x_2 \quad +4x_4+2x_5 \quad = 0 \\ \quad\quad x_3+2x_4 \quad\quad = 0 \\ \quad\quad\quad\quad\quad x_6 = \dfrac{1}{3} \end{cases}$$

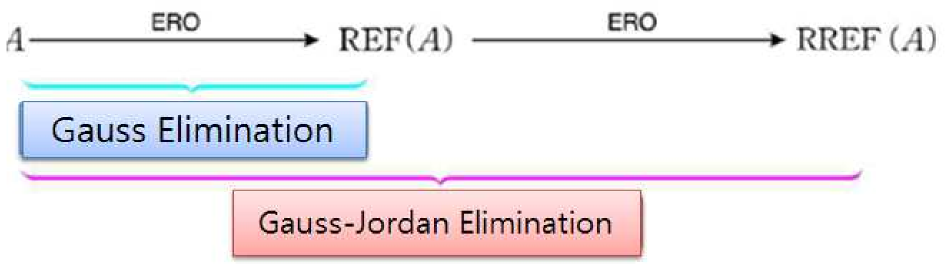
By letting $x_2=r,\ x_4=s,\ x_5=t(r,\ s,\ t$ are any real), its solution is

$$x_1 = -3r - 4s - 2t, \quad x_2 = r, \quad x_3 = -2s, \quad x_4 = s, \quad x_5 = t, \quad x_6 = 1/3. \quad \blacksquare$$

From Example 8, we can express a general solution as a vector form. $(r, s, t \in R)$

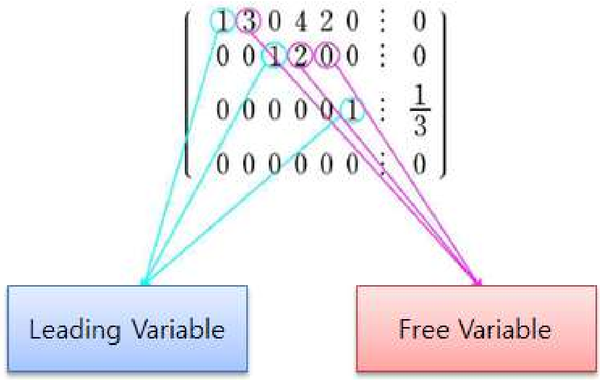
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} -3r-4s-2t \\ r \\ -2s \\ s \\ t \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{3} \end{bmatrix} + \begin{bmatrix} -3r \\ r \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -4s \\ 0 \\ -2s \\ s \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -2t \\ 0 \\ 0 \\ 0 \\ t \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{3} \end{bmatrix} + r\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

[Remark] Gauss elimination and Gauss-Jordan elimination



[Remark] **Leading Variable, Free Variable and their Relation to RREF**

- **a free variable:** a variable corresponding to the column not containing a leading entry in RREF
- **a leading (pivot) variable:** a variable corresponding to the column containing a leading entry in RREF



Homogeneous linear system

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0
\end{aligned} \qquad \text{(II)}
$$

It is easy to see that $\mathbf{x} = \mathbf{0} = (0, 0, \ldots, 0)$ is always a solution of a homogeneous system (II). $\mathbf{x} = \mathbf{0}$ is called a trivial solution. Also if $\mathbf{x}$ is a solution of (II) then any scalar multiple $c\,\mathbf{x} = \mathbf{0}$ is also a solution of (II). Similarly if $\mathbf{x}$ and $\mathbf{y}$ are two solutions of a homogeneous system, so is their sum. This shows that any homogeneous system has either a trivial solution or infinitely many solutions.

Example **9** Using the Gauss-Jordan elimination, express the solution of the following homogeneous equation as a vector form.

$$
\begin{cases}
x_1 + 3x_2 - 2x_3 + 2x_5 = 0 \\
2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 = 0 \\
5x_3 + 10x_4 + 15x_6 = 0 \\
2x_1 + 6x_2 + 8x_4 + 4x_5 + 18x_6 = 0
\end{cases}
$$

Solution Its augmented matrix is $\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & \vdots & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 & \vdots & 0 \\ 0 & 0 & 5 & 10 & 0 & 15 & \vdots & 0 \\ 2 & 6 & 0 & 8 & 4 & 18 & \vdots & 0 \end{bmatrix}$, and its RREF is $\begin{bmatrix} 1 & 3 & 0 & 4 & 2 & 0 & \vdots & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & \vdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \vdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \vdots & 0 \end{bmatrix}$. Thus, leading entry 1's correspond to leading variables x_1, x_3, x_6 and the rest variables x_2, x_4, x_5 to free variables. We have the following.

$$x_1 = -3x_2 - 4x_4 - 2x_5,\ x_3 = -2x_4,\ x_6 = 0$$

Now let free variables be $x_2 = r,\ x_4 = s,\ x_5 = t,$ then

$$x_1 = -3r - 4s - 2t,\ x_2 = r,\ x_3 = -2s,\ x_4 = s,\ x_5 = t,\ x_6 = 0.$$

Therefore $$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = r\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$ ■

Theorem 2.2.2 [No. of free variables in a homogeneous linear system]

In a homogeneous linear system with n unknowns, if the RREF of the augmented matrix has k leading 1's, the solution set has $n-k$ free variables.

Theorem 2.2.3

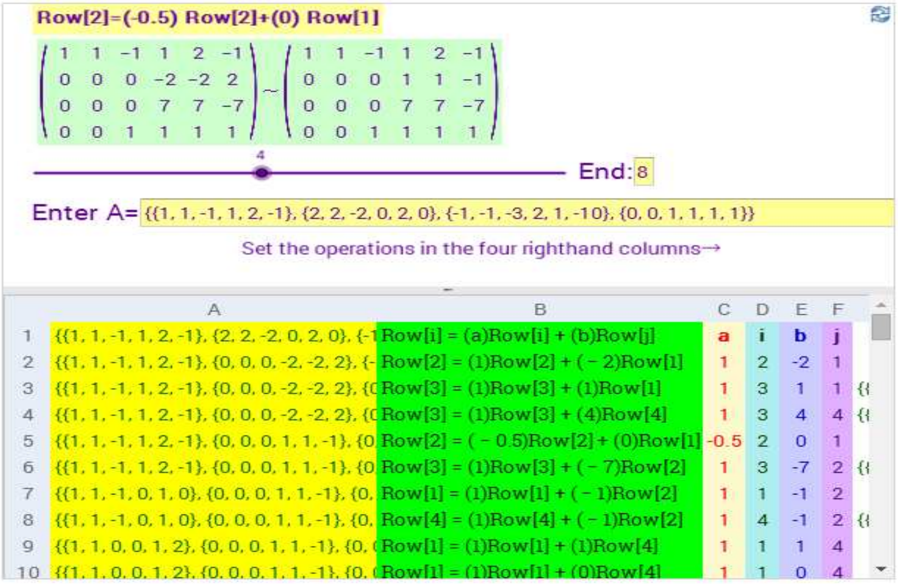
The system $\sum_{j=1}^{n} a_{ij}x_j = 0$ for $1 \le i \le m$ always has a non-trivial solution if $m < n$.

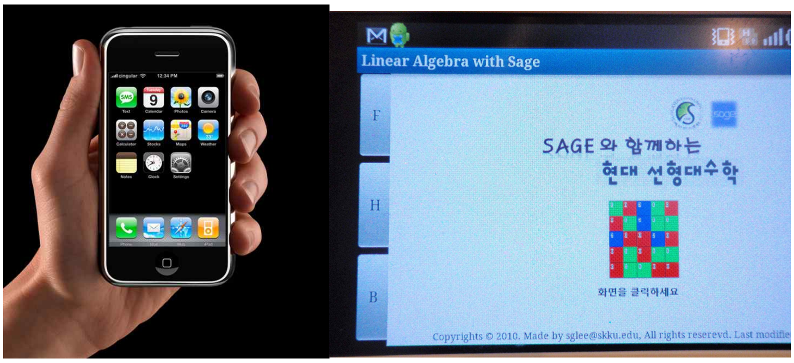
The theorem can be proved using induction on the number on variables.

[Remark] Computer simulation

[Elementary row operation]

http://www.geogebratube.org/student/b73259#material/28831





[Linear algebra with Sage, Smartphone App]

https://play.google.com/store/apps/details?id=la.sage

Chapter 2 Exercises

- http://matrix.skku.ac.kr/LA-Lab/index.htm
- http://matrix.skku.ac.kr/knou-knowls/cla-sage-reference.htm

Problem 1 Answer the questions for the following linear system.

$$\begin{aligned} x_1 + 3x_2 - x_3 &= 1 \\ 2x_1 + 5x_2 + x_3 &= 5 \\ x_1 + \ \ x_2 + x_3 &= 3 \end{aligned}$$

(1) Find the coefficient matrix.

(2) Express the linear system in the form $A\mathbf{x} = \mathbf{b}$.

(3) Find its augmented matrix.

Problem 2 Find a linear system with its augmented matrix.
(Put the unknowns as $x_1, \cdots, x_n$.)

$$\left[\begin{array}{cccc|c} 1 & 1 & 3 & -3 & 0 \\ 0 & 2 & 1 & -3 & 3 \\ 1 & 0 & 2 & -1 & -1 \end{array}\right]$$

Problem 3 Find the number of leading variables and free variables in the solution set of the following system.

$$\begin{cases} x_1 + 4x_2 + 5x_3 - 9x_4 - 7x_5 = 1 \\ \qquad 2x_2 + 4x_3 - 6x_4 - 6x_5 = 2 \\ \qquad\qquad\qquad -5x_4 \qquad\quad = 3 \end{cases}$$

Problem 4 Which matrices are REF or RREF? If one is not RREF, transform it to RREF.

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

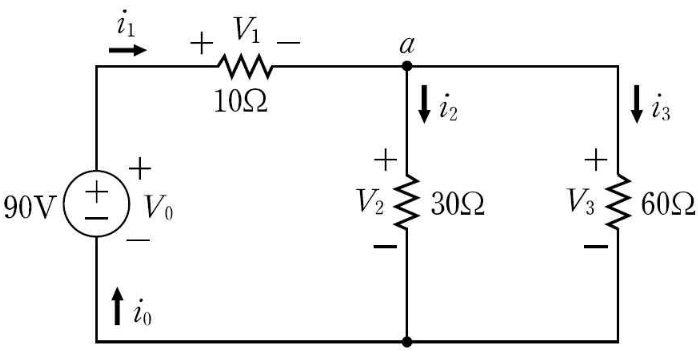
Problem 5 Solve the system using Gauss elimination.

$$\begin{aligned} 2x + y + z - 2w &= 1 \\ 3x - 2y + z - 6w &= -2 \\ x + y - z - w &= -1 \\ 5x - y + 2z - 8w &= 3 \end{aligned} \quad .$$

Problem 6 Solve the system using Gauss-Jordan elimination.

$$\begin{aligned} x + 2y - 3z &= 4 \\ x + 3y + z &= 11 \\ 2x + 5y - 4z &= 13 \\ 2x + 6y + 2z &= 22 \end{aligned} \quad .$$

Problem 7 In the following circuit, write a linear system to find current I_i.



Solution Let $i_0 - i_1 = 0$, $i_0 - i_2 - i_3 = 0$, $i_1 + 3i_2 = 9$, $i_2 - 2i_3 = 0$.

Then $A\mathrm{x} = \mathrm{b}$ where $A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & -1 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & -2 \end{bmatrix}$, $\mathrm{x} = \begin{bmatrix} i_0 \\ i_1 \\ i_2 \\ i_3 \end{bmatrix}$ and $\mathrm{b} = \begin{bmatrix} 0 \\ 0 \\ 9 \\ 0 \end{bmatrix}$. ■

Problem PI In general, we are given a linear system with m equations and n unknowns.

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

If there are k free variables, what is the number of leading variables? From this, think about the relation among the numbers of free variables, leading variables, and unknowns.

Problem p2 Write a linear system with 4 unknowns and 3 equations whose solution set is given below.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + s\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -3 \\ 0 \\ 2 \\ 1 \end{bmatrix} \quad \text{(here } s,\ t \text{ are any real)}$$

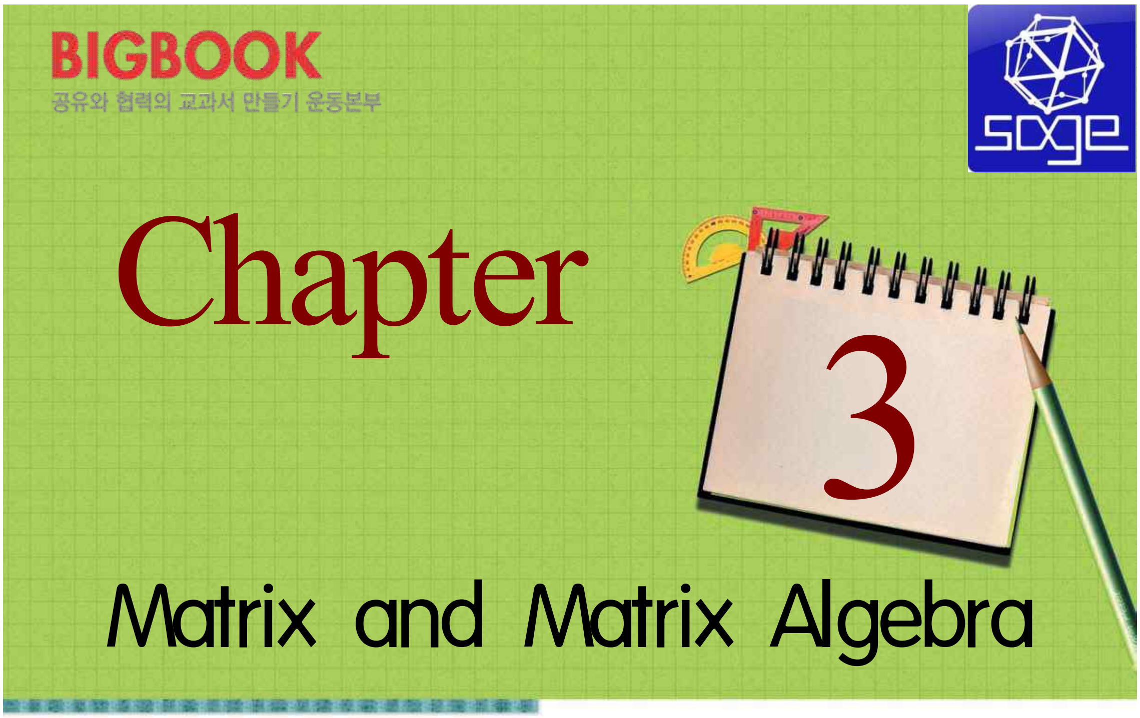
Solution The linear system $\begin{aligned} &x_1 + 2x_2 + 3x_4 - 1 = 0 \\ &x_3 - 2x_4 - 1 = 0 \\ &x_1 + 2x_2 - x_3 + 5x_4 = 0 \end{aligned}$ is an example. ■

Problem p3 Suppose that three points $(1, 4)$, $(-1, 6)$, $(2, 9)$ pass through the parabola $ax^2 + bx + c = y$. By plugging in these points, obtain three linear equations. Find coefficients a, b, c by solving $A\mathrm{x} = \mathrm{b}$.

Problem p4 Write a linear system with four unknowns and four equations satisfying each condition below.

(a) A solution set with one unknown.

(b) A solution set with two unknowns.



Matrix is widely used as a tool to transmit digital sounds and images through internet as well as solving linear systems. We define the addition and product of two matrices. These operations are tools to solve various linear systems. **Matrix product also becomes an excellent tool in dealing with function composition.**

In the previous chapter, we have found the solution set using the Gauss elimination method. In this chapter, **we define the addition and scalar multiplication of matrices and introduce algebraic properties of matrix operations.** Then using the Gauss elimination, we show how to **find the inverse matrix.**

Furthermore, **we investigate the concepts such as linearly independence and subspace which are necessary in understanding the structure of a linear system. Then we** describe the relation between solution set and matrix, and **special matrices.**

3.1 Matrix operation

- Reference video: http://youtu.be/DmtMvQR7cwA, http://youtu.be/JdNnHGdJBrQ
- ractice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-3-Sec-3-1.html

This chapter introduces the definition of the addition and scalar multiplication of matrices and the algebraic properties of the matrix operations. Although many of the properties are identical to those of the operations on real numbers, some properties are different. Matrix operation is a generalization of the operation on real numbers.

Definition [Equality of Matrices]

Two matrices $A=[a_{ij}]_{m\times n}$ and $B=[b_{ij}]_{m\times n}$ of same size are equal if $a_{ij}=b_{ij}$ for all i, j, and denote it by $A=B$.

- To define equal matrices, the size of two matrices should be the same.

Example 1 For what values of x, y, z, w the two matrices

$$A=\begin{bmatrix}1 & 2 & w\\ 2 & -3 & 4\\ 0 & -4 & 5\end{bmatrix},\ B=\begin{bmatrix}1 & 2 & -1\\ 2 & x & 4\\ y & -4 & z\end{bmatrix}$$

are equal?

Solution For $A=B$, each entry should be equal. Thus (that is, $a_{ij}=b_{ij}$) $w=-1$, $x=-3$, $y=0$, $z=5$. ■

Definition [Addition and scalar multiplication of matrix]

Given two matrices $A=[a_{ij}]_{m\times n}$ and $B=[b_{ij}]_{m\times n}$ and a real number k, the sum $A+B$ of A and B, and the scalar multiple kA of A by k are defined by

$$A+B=[a_{ij}+b_{ij}]_{m\times n},\ kA=[ka_{ij}]_{m\times n}.$$

- To define addition, the size of two matrices should be the same.

Example 2 For $A=\begin{bmatrix}1 & 2 & -4\\ -2 & 1 & 3\end{bmatrix}$, $B=\begin{bmatrix}0 & 1 & 4\\ -1 & 3 & 1\end{bmatrix}$, $C=\begin{bmatrix}1 & 1\\ 2 & 2\end{bmatrix}$, what is $A+B$, $2A$, $(-1)C$?

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

$$2A=\begin{bmatrix}2\cdot 1 & 2\cdot 2 & 2\cdot(-4)\\ 2\cdot(-2) & 2\cdot 1 & 2\cdot 3\end{bmatrix}=\begin{bmatrix}2 & 4 & -8\\ -4 & 2 & 6\end{bmatrix}$$

$$(-1)C=\begin{bmatrix}(-1)\cdot 1 & (-1)\cdot 1\\ (-1)\cdot 2 & (-1)\cdot 2\end{bmatrix}=\begin{bmatrix}-1 & -1\\ -2 & -2\end{bmatrix}.$$ □

- http://matrix.skku.ac.kr/RPG_English/3-MA-operation.html
- http://matrix.skku.ac.kr/RPG_English/3-MA-operation-1.html

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ,[[1,2,-4], [-2,1,3]])
B=matrix(QQ,[[0,1,4], [-1,3,1]])
C=matrix(QQ,[[1,1],[2,2]])
print A+B                    # matrix addition
print
print 2*A                    # scalar multiplication
print
print (-1)*C                 # scalar multiplication
```

```
[ 1  3  0]        [ 2  4 -8]        [-1 -1]
[-3  4  4]        [-4  2  6]        [-2 -2]
```
■

Definition **[Matrix product]**

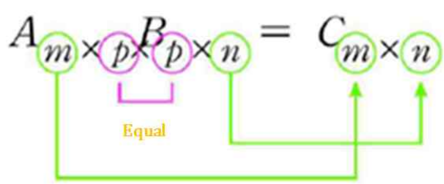
Given two matrices $A=[a_{ij}]_{m\times p}$ and $B=[b_{ij}]_{p\times n}$, the product AB of A and B is defined below.

$$AB=[c_{ij}\ ]_{m\times n},$$

where $c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{ip}b_{pj}\ =\sum_{k=1}^{p}a_{ik}b_{kj}\ (1\le\ i\le\ m,\ 1\le\ j\le\ n)$.

For two matrices A and B to be compatible for multiplication, we require the number of columns of A to be equal to the number of rows of B. The resultant matrix AB is of size number of rows of A by the number of columns of B.

[Remark]



Determining the size of the matrix product

[Remark] **Meaning of matrix product**

Let $A=[a_{ij}]_{m\times p}$, $B=[b_{ij}]_{p\times n}$, and denote the ith row of A by $A_{(i)}$ and the jth column of A by $A^{(j)}$. Then

$$C\ =AB\ =\begin{bmatrix}A_{(1)}\\A_{(2)}\\\vdots\\A_{(m)}\end{bmatrix}\left[B^{(1)}\ B^{(2)}\cdots\ B^{(n)}\right]=\begin{bmatrix}A_{(1)}B^{(1)} & A_{(1)}B^{(2)}\cdots & A_{(1)}B^{(n)}\\ \vdots & & \vdots\\ A_{(m)}B^{(1)} & \cdots & A_{(m)}B^{(n)}\end{bmatrix}_{m\times n}$$

Thus, $c_{ij}=A_{(i)}B^{(j)}=\left[a_{i1}\ \cdots\ a_{ip}\right]\begin{bmatrix}b_{1j}\\\vdots\\b_{pj}\end{bmatrix}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{ip}b_{pj}=\sum_{k=1}^{p}a_{ik}b_{kj}$

Note that the inner product of ith row vector of A and the jth column vector of B is the (i,j) entry of AB.

$$AB=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1p}\\ \vdots & \vdots & & \vdots\\ a_{i1} & a_{i2} & \cdots & a_{ip}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mp}\end{bmatrix}\begin{bmatrix}b_{11} & \cdots & b_{1j} & \cdots & b_{1n}\\ b_{21} & \cdots & b_{2j} & \cdots & b_{2n}\\ \vdots & & \vdots & & \vdots\\ b_{p1} & \cdots & b_{pj} & \cdots & b_{pn}\end{bmatrix}$$

Example 3

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

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

$$=\begin{bmatrix}(1)(-2)+(2)(0)+(-1)(2) & (1)(1)+(2)(-3)+(-1)(1)\\(3)(-2)+(1)(0)+(0)(2) & (3)(1)+(1)(-3)+(0)(1)\end{bmatrix}=\begin{bmatrix}-4&-6\\-6&0\end{bmatrix} \quad \square$$

- http://matrix.skku.ac.kr/RPG_English/3-MA-operation-1-multiply.html

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ,[[1,2,-1], [3,1,0]])
B=matrix(QQ,[[-2,1], [0,-3], [2,1]])
print A*B                          # Don't forget to include (*)!
```

```
[-4 -6]
[-6  0]
```

■

Using matrix product, one can express a linear system easily. Let us consider the following linear system

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

and let $A=[a_{ij}]_{m\times n}$, $\mathbf{x}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$, $\mathbf{b}=\begin{bmatrix}b_1\\b_2\\\vdots\\b_m\end{bmatrix}$ be the coefficient matrix, the unknown vector and the constant vector respectively. Then we can express the linear system as

$$A\mathbf{x}=\mathbf{b} \quad \Leftrightarrow \quad x_1A^{(1)}+x_2A^{(2)}+\cdots+x_nA^{(n)}=\mathbf{b}$$

$$\Leftrightarrow \quad x_1\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{mn} \end{bmatrix}$$

Theorem 3.1.1

Let A, B, C be matrices of proper sizes (oeprations are well defined) and let a, b be scalars. Then the following hold.

(1) $A+B=B+A$ (commutative law of addition)
(2) $A+(B+C)=(A+B)+C$ (associative law of addition)
(3) $A(BC)=(AB)C$ (associative law of multiplication)
(4) $A(B+C)=AB+AC$ (distributive law)
(5) $(B+C)A=BA+CA$ (distributive law)
(6) $a(B+C)=aB+aC$
(7) $(a+b)C=aC+bC$
(8) $(ab)C=a(bC)$
(9) $a(BC)=(aB)C=B(aC)$

The proof of the above facts are easy and readers are encouraged to prove them.

Example 4 Check the associative law of the matrix product.

$$A=\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix},\ B=\begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix},\ C=\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$$

Solution Since $AB=\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}=\begin{bmatrix} 8 & 5 \\ 20 & 13 \\ 2 & 1 \end{bmatrix}$, we have

$$(AB)C=\begin{bmatrix} 8 & 5 \\ 20 & 13 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}=\begin{bmatrix} 18 & 15 \\ 46 & 39 \\ 4 & 3 \end{bmatrix}$$

Since $BC=\begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}=\begin{bmatrix} 10 & 9 \\ 4 & 3 \end{bmatrix}$, we have

$A(BC)=\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 10 & 9 \\ 4 & 3 \end{bmatrix}=\begin{bmatrix} 18 & 15 \\ 46 & 39 \\ 4 & 3 \end{bmatrix}$. Hence, $(AB)C=A(BC)$. ■

- The properties of operations on matrices are similar to those of operations on real numbers which are well known,
- **Exception:** For matrices A, B, we do not have $AB=BA$ in general.

Example 5 Suppose that we are given the following matrices A, B, C, D, E.

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

$$D=\begin{bmatrix}-1 & 0\\ 2 & 3\end{bmatrix},\quad E=\begin{bmatrix}1 & 2\\ 3 & 0\end{bmatrix}.$$

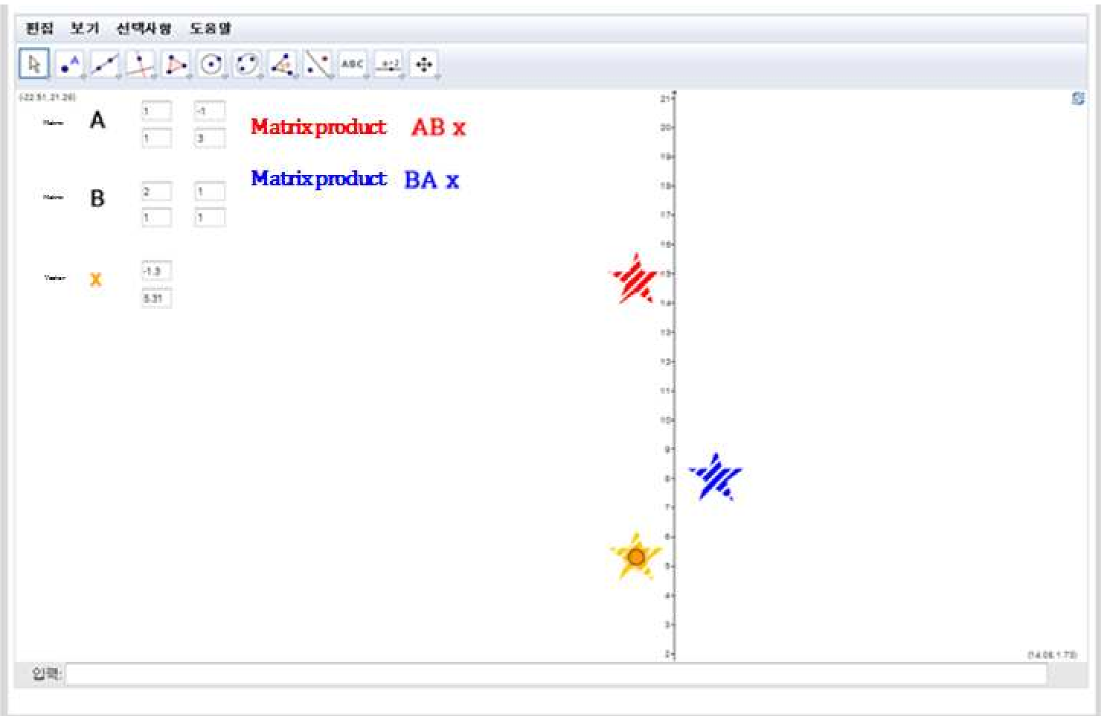
Then AB is defined but BA is not defined. Similarly AC is a 2×2 matrix but CA is a 3×3 matrix, and hence $AC\neq CA$. Also although DE and ED are 2×2 matrices, as we can see below, we have $DE\neq ED$.

$$DE=\begin{bmatrix}-1 & -2\\ 11 & 4\end{bmatrix},\quad ED=\begin{bmatrix}3 & 6\\ -3 & 0\end{bmatrix}$$ ■

[Remark] **Computer simulation**

[matrix product] (Commutative law does not hold.)

http://www.geogebratube.org/student/m12831



Definition **[Zero matrix]**

A zero matrix consists of entries of 0's and denoted by O(or $O_{m \times n}$).

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, [0] \cdots$$

Theorem 3.1.2

For any matrix A and a zero matrix O of a proper size, the following hold.

(1) $A+O=O+A=A$

(2) $A-A=O$

(3) $O-A=-A$

(4) $AO=OA=O$

- **Note:** Although $AB=O$, it is possible to have $A \neq O$, $B \neq O$. Similarly, although $AB=AC$, $A \neq O$, it is possible to have $B \neq C$.

Example 6 Let $A=\begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}$, $B=\begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix}$, $C=\begin{bmatrix} 2 & 5 \\ 3 & 4 \end{bmatrix}$, $D=\begin{bmatrix} 3 & 7 \\ 0 & 0 \end{bmatrix}$. Then $AB=\begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}=AC$

But $A \neq O$ and $B \neq C$. Also $AD=O$ but $A \neq O$, $D \neq O$. ■

We should first define scalar matrices.

Definition **[Identity matrix]**

A scalar matrix of order n with diagonal entries all 1's is called an **identity matrix** of order n and is denoted by I_n. That is,

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

- Let A be an $m \times n$ matrix and the identity matrix I_m, I_n. It is easy to see that

$$I_m A = A = A I_n.$$

Example 7

Let $A=\begin{bmatrix}4&-2&3\\5&0&2\end{bmatrix}$. Then $I_2A=\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}4&-2&3\\5&0&2\end{bmatrix}=\begin{bmatrix}4&-2&3\\5&0&2\end{bmatrix}=A$,

$AI_3=\begin{bmatrix}4&-2&3\\5&0&2\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}=\begin{bmatrix}4&-2&3\\5&0&2\end{bmatrix}=A$.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ,[[4,-2,3], [5,0,2]])
I2=identity_matrix(2)       # identity matrix identity_matrix(n), n is the order
I3=identity_matrix(3)
O2=zero_matrix(3, 2)        # zero matrix zero_matrix(m, n), m, n are the order
print I2*A
print
print A*I3
print
print A*O2
```

```
[ 4 -2  3]
[ 5  0  2]

[ 4 -2  3]
[ 5  0  2]

[0 0]
[0 0]
```

■

Definition

Let A be a square matrix of order n. The kth **power** of A is defined by

$$A^0=I_n,\ A^k=AA\cdots A(k \text{ times})$$

Theorem 3.1.3

If A is a square matrix and r, s are non negative integers, then

$$A^rA^s=A^{r+s},\quad (A^r)^s=A^{rs}.$$

Example 8 Let $A = \begin{bmatrix} 4 & -2 \\ 5 & 0 \end{bmatrix}$. Find A^2, A^3, A^0 and confirm that $(A^2)^3 = A^6$.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ,[[4,-2], [5,0]])
print A^2                  # Works only for a square matrix
print
print A^3                  # same format as power of real numbers
print
print A^0                  # When the exponent is 0, get identity matrix
print
(A^2)^3==A^6               # check the power rule
```

```
[  6  -8]
[ 20 -10]

[-16 -12]
[ 30 -40]

[1 0]
[0 1]

True
```

■

In the set of real numbers, we have $(a+b)^2 = a^2 + ab + ba + b^2 = a^2 + ab + ab + b^2 = a^2 + 2ab + b^2$. However, the commutative law under matrix product does not work and thus we only have the following.

$$(A+B)^2 = A^2 + AB + BA + B^2.$$

When $AB = BA$, we have $(A+B)^2 = A^2 + 2AB + B^2$.

Definition [Transpose matrix]

For a matrix $A = [a_{ij}]_{m \times n}$, the **transpose** of A is denoted by A^T and defined by

$$A^T = [a_{ij}']_{n \times m},\ a_{ij}' = a_{ji} \quad (1 \le i \le n, 1 \le j \le m).$$

- The transpose A^T of A is obtained by interchanging the rows and columns of A.

Example 9

Find the transpose of the following matrices.

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

Solution $$A^T=\begin{bmatrix}1&4\\-2&5\\3&0\end{bmatrix},\ B^T=\begin{bmatrix}1&3&0\\2&-1&5\\-4&2&3\end{bmatrix},\ C^T=\begin{bmatrix}5&-3&2\\4&2&1\end{bmatrix},$$

$$D^T=\begin{bmatrix}3\\0\\1\end{bmatrix},\ E^T=[2\ \ 0\ \ -3].$$ □

Sage http://sage.skku.edu 또는 http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ,[[1,-2,3], [4,5,0]])
C=matrix(QQ,[[5,4], [-3,2], [2,1]])
D=matrix(QQ, [[3,0,1]])
print A.transpose()          # Transpose of a matrix  A.transpose()
print
print C.transpose()
print
print D.transpose()
```

```
[ 1  4]        [ 5 -3  2]        [3]
[-2  5]        [ 4  2  1]        [0]
[ 3  0]                          [1]
```
■

Theorem 3.1.4

Let A, B be matrices of appropriate sizes and k a scalar. The following hold.

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

(2) $(A+B)^T = A^T + B^T$

(3) $(AB)^T = B^T A^T$

(4) $(kA)^T = kA^T$.

Example 10

Let $A=\begin{bmatrix}1&3\\2&5\end{bmatrix}$, $B=\begin{bmatrix}1&1&3\\2&1&4\end{bmatrix}$. Show that (3) of Theorem 3.1.4 is true.

Solution Since $AB=\begin{bmatrix}1&3\\2&5\end{bmatrix}\begin{bmatrix}1&1&3\\2&1&4\end{bmatrix}=\begin{bmatrix}7&4&15\\12&7&26\end{bmatrix}$, $(AB)^T=\begin{bmatrix}7&12\\4&7\\15&26\end{bmatrix}$.

Also, $B^TA^T=\begin{bmatrix}1&2\\1&1\\3&4\end{bmatrix}\begin{bmatrix}1&2\\3&5\end{bmatrix}=\begin{bmatrix}7&12\\4&7\\15&26\end{bmatrix}$. Thus $(AB)^T=B^TA^T$. ■

Definition [Trace]

The **trace** of $A=[a_{ij}]_{n\times n}$ is defined by $\mathrm{tr}(A)=a_{11}+a_{22}+\cdots+a_{nn}=\sum_{i=1}^{n}a_{ii}$.

Theorem 3.1.5

If A, B are square matrices of the same size and $c\in R$, then

(1) $\mathrm{tr}(A^T)=\mathrm{tr}(A)$

(2) $\mathrm{tr}(cA)=c\,\mathrm{tr}(A)$, $c\in R$

(3) $\mathrm{tr}(A+B)=\mathrm{tr}(A)+\mathrm{tr}(B)$

(4) $\mathrm{tr}(A-B)=\mathrm{tr}(A)-\mathrm{tr}(B)$

(5) $\mathrm{tr}(AB)=\mathrm{tr}(BA)$

Proof We prove the item (5) only and leave the rest as an exercise.

$$\mathrm{tr}(AB)=\sum_{i=1}^{n}\left(\sum_{k=1}^{n}a_{ik}b_{ki}\right)=\sum_{k=1}^{n}\sum_{i=1}^{n}b_{ki}a_{ik}=\sum_{k=1}^{n}\left(\sum_{i=1}^{n}b_{ki}a_{ik}\right)=\mathrm{tr}(BA).$$ ■

Example 11

Let $A=\begin{bmatrix}1&-2\\4&5\end{bmatrix}$, $B=\begin{bmatrix}5&4\\-3&2\end{bmatrix}$. Show that (5) of Theorem 3.1.5 is true.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ,[[1,-2], [4,5]])
B=matrix(QQ,[[5,4], [-3,2]])
print (A*B).trace()             # trace. A.trace()
print
print (B*A).trace()
```

```
37

37
```
■

3.2 Inverse matrix

- Reference video: http://youtu.be/GCKM2VlU7bw, http://youtu.be/yeCUPdRx7Bk
- Practice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-3-Sec-3-2.html

In this chapter, we introduce an inverse matrix of a square matrix which plays like a multiplicative inverse of a real number. We investigate the properties of an inverse matrix. You will see that some properties holding in the inverse of a real number are not true in the matrix inverse operation although most hold in both inverses.

Definition

A square matrix A of order n is called **invertible** (or **nonsingular**) if there is a square matrix B such that

$$AB = I_n = BA.$$

This matrix B if exists is called the **inverse matrix** of A. If such a matrix B does not exist, A is called **noninvertible**, (or **singular**).

Example 1

From matrices $A = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$, we see that B is the inverse matrix of A by the following computation.

$$AB = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2,$$

$$BA = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$ ■

Example 2

Let $A = \begin{bmatrix} 1 & 4 & 3 \\ 2 & 5 & 6 \\ 0 & 0 & 0 \end{bmatrix}$. Note that the third row of A has all zeroes. Thus for any matrix

$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$ the third row of AB is $[0 \;\; 0 \;\; 0]$. Therefore there does not exist B such that $AB = I$, that is, A is singular.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ,[[1,4,3],[2,5,6],[0,0,0]])
A.is_invertible()          # check if matrix is invertible A.is_invertible()
```

False ■

Theorem 3.2.1

If A is an invertible square matrix of order n, then an inverse of A is unique.

Proof Suppose that B, C are inverses of A. Then as

$$AB = BA = I_n, \quad AC = CA = I_n$$

we have

$$B = BI_n = B(AC) = (BA)C = I_n C = C$$

Thus an inverse of A is unique. ■

Example 3 A necessary and sufficient condition for $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ to be invertible is that $ad - bc \neq 0$. Hence one has

$$A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}.$$

It is straightforward to check

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$ ■

Theorem 3.2.2

If A, B are invertible square matrices of order n and k is a nonzero scalar, then the following hold.

(1) A^{-1} is invertible and $(A^{-1})^{-1} = A$.

(2) AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

(3) kA is invertible and $(kA)^{-1} = \frac{1}{k}A^{-1}$.

(4) A^n is invertible and $(A^n)^{-1} = A^{-n} = (A^{-1})^n$.

Proof (1)~(4) Just check that the product of matrices are the identity matrix. ■

Theorem 3.2.3

If A is an invertible matrix, then so is A^T and the following holds.

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

Example 4 Let $A = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$. Check that $(AB)^{-1} = B^{-1}A^{-1}$.

Solution Since $A^{-1} = \frac{1}{6-5}\begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}$,

$B^{-1} = \frac{1}{7-6}\begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix}$, we have

$B^{-1}A^{-1} = \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 17 & -44 \\ -5 & 13 \end{bmatrix}$. Also since

$AB = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix} = \begin{bmatrix} 13 & 44 \\ 5 & 17 \end{bmatrix}$ we have

$(AB)^{-1} = \frac{1}{221-220}\begin{bmatrix} 17 & -44 \\ -5 & 13 \end{bmatrix} = \begin{bmatrix} 17 & -44 \\ -5 & 13 \end{bmatrix}$. ■

- http://matrix.skku.ac.kr/RPG_English/3-SO-MA-inverse.html

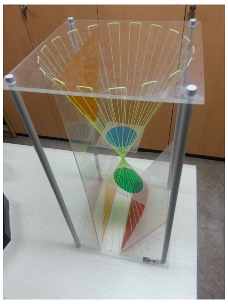
Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

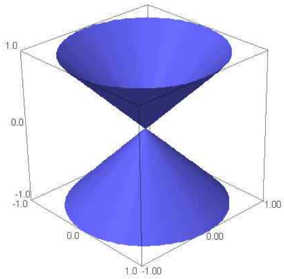
```
A=matrix(ZZ, 2, 2, [3, 5, 1, 2])
B=matrix(ZZ, 2, 2, [1, 3, 2, 7])
AB=A*B                          # AB calculation
print AB.inverse()              # inverse of AB, format A.inverse()
print
print B.inverse()*A.inverse()   # B^(-1)*A^(-1)
```

```
[ 17 -44]
[ -5  13]

[ 17 -44]
[ -5  13]
```

■





<3D printing Object of Conic Section>
http://www.youtube.com/watch?v=q_XPFJjncmQ&feature=youtu.be

3.3 Elementary matrices

- Reference video: http://youtu.be/GCKM2VlU7bw, http://youtu.be/oQ2m6SSSquc
- Practice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-3-Sec-3-3.html

In the previous section, we defined an inverse of square matrices. In this section, we shall discuss how to find an inverse of square matrices by using elementary row operations and elementary matrices.

Definition

An n by n matrix is called an **elementary matrix** if it can be obtained from I_n by performing a single **elementary row operation (ERO).** A **permutation matrix** is obtained by exchanging rows of I_n.

Example 1 Listed below are three elementary matrices and the operations that produce them.

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$: Interchange the 2nd and the 3th rows. $R_2 \leftrightarrow R_3$

$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$: Add 2 times the 1st row to the 2nd row. $2R_1 + R_2 \rightarrow R_2$

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$: Multiply the 2nd row by 3. $3R_2 \rightarrow R_2$

Sage http://sage.skku.edu (Warning!! The index of Sage starts from 0.)

```
E1=elementary_matrix(4, row1=1, row2=2)    # elementary matrix r2 <--> r3
# elementary_matrix(n, row1=i, row2=j)        exchange of ith row, jth row
E2=elementary_matrix(4, row1=2, scale=-3)   # elemenatry matrix (-3)*r3
# elementary_matrix(n, row1=i, scale=m)       multiply ith row by m
E3=elementary_matrix(4, row1=0, row2=3, scale=7) # row 7*r4 + r1
# elementary_matrix(n, row1=i, row2=j, scale=m)    add m times jth row to the ith row.
print E1
```

```
print E2
print E3
```

```
[1 0 0 0]        [ 1   0   0   0]        [1 0 0 7]
[0 0 1 0]        [ 0   1   0   0]        [0 1 0 0]
[0 1 0 0]        [ 0   0  -3   0]        [0 0 1 0]
[0 0 0 1]        [ 0   0   0   1]        [0 0 0 1]
```

■

Example 2 [Property of elementary matrix] The product of an elementary matrix E on the left and any matrix A is the matrix that results when the corresponding same row operation is performed on A.

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 0 & 1 & 3 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 0 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 0 & 1 & 3 \end{bmatrix} \xrightarrow{2R_1 + R_2} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 5 & 7 \\ 0 & 1 & 3 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 0 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 5 & 7 \\ 0 & 1 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 0 & 1 & 3 \end{bmatrix} \xrightarrow{3R_2} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 0 & 1 & 3 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 0 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 0 & 1 & 3 \end{bmatrix}$$

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ, 3,3, [1,2,3,1,1,1,0,1,3])
E1=elementary_matrix(3, row1=1, row2=2)           # r2 <--> r3
E2=elementary_matrix(3, row1=1, row2=0, scale=2)  # 2*r1 + r2
E3=elementary_matrix(3, row1=1, scale=3)          #  3*r2
print E1*A
print
print E2*A
print
print E3*A
```

```
[1 2 3]        [1 2 3]        [1 2 3]
[0 1 3]        [3 5 7]        [3 3 3]
[1 1 1]        [0 1 3]        [0 1 3]
```

■

[Remark] The inverse of an elementary matrix is elementary.

Since $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

Since $\begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/k & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/k & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Since $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & c & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -c & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & c & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -c & 1 \end{bmatrix}$

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
E1=elementary_matrix(3, row1=1, row2=2)            # r2 <--> r3
E2=elementary_matrix(3, row1=2, row2=1, scale=4)   # 4*r2 + r3
E3=elementary_matrix(3, row1=1, scale=3)           # 3*r2
print E1.inverse()
print
print E2.inverse()
print
print E3.inverse()
```

```
[1 0 0]        [ 1  0  0]        [  1   0   0]
[0 0 1]        [ 0  1  0]        [  0 1/3   0]
[0 1 0]        [ 0 -4  1]        [  0   0   1]
```

Finding the inverse of an invertible matrix.

We investigate the method to find the inverse of an invertible matrix using elementary matrices. First consider equivalent statements of an invertible matrix (its proof will be treated in Chapter 7).

Theorem 3.3.1 [Equivalent statements]

For any $n \times n$ matrix A, the followings are equivalent.

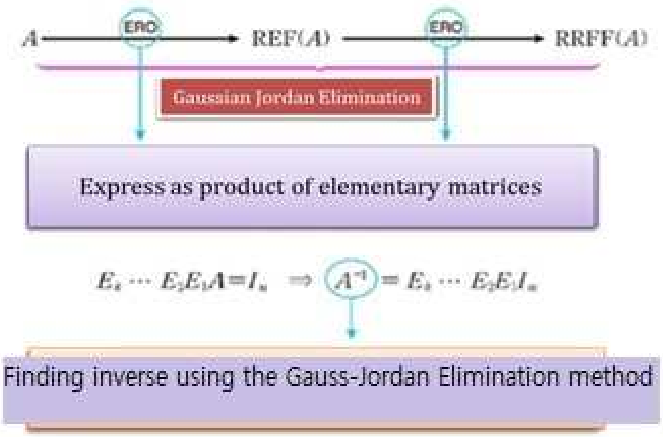
(1) A is invertible.

(2) A is row equivalent to I_n. (i.e. $\text{RREF}(A) = I_n$)

(3) A can be expressed as a product of elementary matrices.

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

[Remark]



Theorem 3.3.2 [Computation of an inverse]

$$[A \ : I_n] \xrightarrow{RREF} [I_n \ : A^{-1}]$$

[Remark] **Finding an inverse using the Gauss-Jordan elimination.**

[Step 1] For a given A, augment I_n on the right side so that we make a $n \times 2n$ matrix $[A \ : \ I_n]$.

[Step 2] Compute the RREF of $[A \vdots \ I_n]$.

[Step 3] Let $[C \vdots \ D]$ be the RREF of $[A \vdots \ I_n]$ in the step 2. Then, following hold.

(i) If $C = I_n$, then $D = A^{-1}$.

(ii) If $C \neq I_n$, then A is not invertible so that A^{-1} does not exist.

Example 3 Find the inverse of

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

Solution Consider $[A \vdots \ I_3]$. Then

$$[A : \ I_3] = \begin{bmatrix} 1 & 2 & 3 & \vdots & 1 & 0 & 0 \\ 2 & 5 & 3 & \vdots & 0 & 1 & 0 \\ 1 & 0 & 8 & \vdots & 0 & 0 & 1 \end{bmatrix}$$

and, its RREF is given as follows.

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

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

$$\rightarrow \begin{bmatrix} 1 & 2 & 0 & \vdots & -14 & 6 & 3 \\ 0 & 1 & 0 & \vdots & 13 & -5 & -3 \\ 0 & 0 & 1 & \vdots & 5 & -2 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & \vdots & -40 & 16 & 9 \\ 0 & 1 & 0 & \vdots & 13 & -5 & -3 \\ 0 & 0 & 1 & \vdots & 5 & -2 & -1 \end{bmatrix} = [C : \ D]$$

Since $C = I_3$, $D = A^{-1}$.

$$\therefore \ A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$$ ■

Example **4** Find the inverse of

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

Solution It follows from a similar way to Example 03,

$$\begin{bmatrix} 1 & 6 & 4 & \vdots & 1 & 0 & 0 \\ 2 & 4 & -1 & \vdots & 0 & 1 & 0 \\ -1 & 2 & 5 & \vdots & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 6 & 4 & \vdots & 1 & 0 & 0 \\ 0 & -8 & -9 & \vdots & -2 & 1 & 0 \\ 0 & 8 & 9 & \vdots & 1 & 0 & 1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 6 & 4 & \vdots & 1 & 0 & 0 \\ 0 & -8 & -9 & \vdots & -2 & 1 & 0 \\ 0 & 0 & 0 & \vdots & -1 & 1 & 1 \end{bmatrix} = [C : \ D]$$

Since $C \neq I_3$, A^{-1} does not exist. $C \neq I_3$. ■

Example **5** Find the inverse of

$$A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 0 & 2 \\ -6 & 4 & 11 \end{bmatrix}$$

- http://matrix.skku.ac.kr/RPG_English/3-MA-Inverse_by_RREF.html

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ, 3, 3, [1, -1, 2, -1, 0, 2, -6, 4, 11])
I=identity_matrix(3)
Aug=A.augment(I).echelon_form() # augmented matrix [A : I] echelon_form
show(Aug)
```

```
[    1     0     0  |  8/15 -19/15   2/15]
[    0     1     0  |  1/15 -23/15   4/15]
[    0     0     1  |  4/15 -2/15    1/15]
```

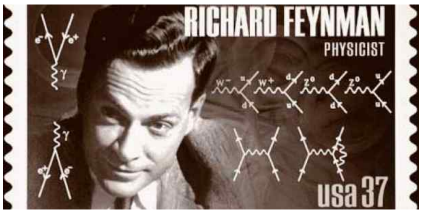
We can extract inverse of A using slicing of the above matrix.

```
Aug[:, 3:6]
```

```
[  8/15 -19/15   2/15]
[  1/15 -23/15   4/15]
[  4/15  -2/15   1/15]
```

Thus $A^{-1} = \frac{1}{15}\begin{bmatrix} 8 & -19 & 2 \\ 1 & -23 & 4 \\ 4 & -2 & 1 \end{bmatrix}$. ■

'If you want to learn about nature, to appreciate nature, it is necessary to understand the language(Mathematics) that she speaks in.'



Richard Phillips Feynman (1918-1988) was an American theoretical physicist known for his work in the path integral formulation of quantum mechanics, the theory of quantum electrodynamics, and the physics of the superfluidity of supercooled liquid helium, as well as in particle physics.

3.4 Subspaces and Linear Independence

- Reference video: http://youtu.be/HFq_-8B47xM. http://youtu.be/UTTUg6JUFQM
- Practice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-4-Sec-3-4.html

In this section, we define a linear combination, a spanning set, a linear (in)dependence and a subspace of $\mathbb{R}^n$. We will also learn how to solve the system of linear equations by using the fact that solutions for a system of homogeneous linear equations form a subspace of $\mathbb{R}^n$.

Note that $\mathbb{R}^n$ with standard addition and scalar multiplication is also called a vector space over $\mathbb{R}$ and its elements are called vectors.

Definition [Subspace]

Let W be a nonempty subset of $\mathbb{R}^n$. Then W is called a **subspace** of $\mathbb{R}^n$ if W satisfies the following two conditions.

(1) $\mathbf{x},\ \mathbf{y} \in W \ \Rightarrow\ \mathbf{x}+\mathbf{y} \in W$ (closed under the addition)
(2) $\mathbf{x} \in W,\ k \in \mathbb{R} \ \Rightarrow\ k\mathbf{x} \in W$ (closed under the scalar multiplication)

- All subspaces of $\mathbb{R}^n$ contain zero vector.
  $\mathbf{x} \in W,\ 0 \in \mathbb{R} \ \Rightarrow\ 0\mathbf{x} = \mathbf{0} \in W$

Example 1 $\{\mathbf{0}\}$ and $\mathbb{R}^n$ are subspaces of $\mathbb{R}^n$ where $\mathbf{0} = (0, 0, \cdots, 0)$ is denoted by the origin. They are called the trivial subspaces. ■

Example 2 A subset $L_0 = \{(x,y) \in \mathbb{R}^2 \mid y = x\}$ of $\mathbb{R}^2$ satisfies two conditions for subspace. Hence, L_0 is a subspace of $\mathbb{R}^2$. On the other hand, a subset $L_1 = \{(x,y) \in \mathbb{R}^2 \mid y = x+1\}$ of $\mathbb{R}^2$ does not satisfy conditions for subspace so that L_1 is not a subspace of $\mathbb{R}^2$.

$(0,1),\ (1,2) \in L_1$ but $(0,1)+(1,2) = (1,3) \notin L_1$ ■

Example 3 All subspaces of $\mathbb{R}^2$ are one of the followings.

1. zero subspace : $\{\mathbf{0}\}$
2. Lines through the origin.
3. $\mathbb{R}^2$

All subspaces of $\mathbb{R}^3$ are one of the following.

1. zero subspace : $\{\mathbf{0}\}$
2. Lines through the origin
3. Planes through the origin
4. $\mathbb{R}^3$ ■

■

Example 4 Show that a subset $W = \{(0, a, b, c, d, 0) \,|\, a, b, c, d \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^6$.

Solution For

$$\mathbf{x} = (0, a_1, b_1, c_1, d_1, 0),\ \mathbf{y} = (0, a_2, b_2, c_2, d_2, 0) \in W,\ k \in \mathbb{R}$$

the following hold.

(i) $\mathbf{x}+\mathbf{y} = (0, a_1+a_2, b_1+b_2, c_1+c_2, d_1+d_2, 0) \in W$

(ii) $k\mathbf{x} = (0, ka_1, kb_1, kc_1, kd_1, 0) \in W$

Therefore, W is a subspace of $\mathbb{R}^6$. ■

Let $M_{m\times n}$ denote the set of all $m\times n$ matrices over $\mathbb{R}$.

Example 5 For $A \in M_{m\times n}$, show that

$$W = \{\mathbf{x} \in \mathbb{R}^n \,|\, A\mathbf{x} = \mathbf{0}\}$$

is a subspace of $\mathbb{R}^n$. (This W is called a **solution space** or **null space** of A)

Solution Clearly, $A\mathbf{0} = \mathbf{0}$ so that $\mathbf{0} \in W$, $W \neq \varnothing$. Since for $\mathbf{x}, \mathbf{y} \in W$, $k \in \mathbb{R}$

$$A\mathbf{x} = \mathbf{0},\ A\mathbf{y} = \mathbf{0},$$

we can obtain that

$$A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0}+\mathbf{0} = \mathbf{0} \text{ and}$$

$$A(k\mathbf{x}) = k(A\mathbf{x}) = k\mathbf{0} = \mathbf{0}.$$

This implies $\mathbf{x}+\mathbf{y} \in W,\ k\mathbf{x} \in W$.

Therefore, W is a subspace of $\mathbb{R}^n$. ■

Definition **[linear combination]**

If $\mathbf{x} \in \mathbb{R}^n$ can be expressed in the form

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k, \quad c_1, c_2, \ldots, c_k \in \mathbb{R}$$

with $\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k\} \subseteq \mathbb{R}^n$, then $\mathbf{x}$ is called a **linear combination** of vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$.

Example 6 Let $\mathbf{x}_1 = (1, -2, -1)$, $\mathbf{x}_2 = (3, -5, 4)$ be vectors of $\mathbb{R}^3$. Can $\mathbf{x} = (2, -6, 3)$ be a linear combination of $\mathbf{x}_1$ and $\mathbf{x}_2$?

Solution The answer is depend on whether there exist c_1, c_2 in $\mathbb{R}$ such that

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2.$$

From this observation, we can obtain

$$\begin{bmatrix} 2 \\ -6 \\ 3 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} + c_2 \begin{bmatrix} 3 \\ -5 \\ 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 3 \\ -2 & -5 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -6 \\ 3 \end{bmatrix} \qquad \square$$

One can easily show that the above system has no solution.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(3, 3, [1, 3, 2, -2, -5, -6, -1, 4, 3])  # augmented matrix
print A.rref()
```

```
[1 0 0]
[0 1 0]
[0 0 1]
```

Since this system of linear equation has no solution, there are no such scalars c_1, c_2 exist. Consequently, $\mathbf{x}$ is not a linear combination of $\mathbf{x}_1, \mathbf{x}_2$. ■

Example 7 Show that the set of all linear combinations of $S=\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k\} \subseteq \mathbb{R}^n$

$$W=\{c_1\mathbf{x}_1+c_2\mathbf{x}_2+\cdots+c_k\mathbf{x}_k \mid c_1, c_2, \cdots, c_k \in \mathbb{R}\}$$

is a subspace of $\mathbb{R}^n$.

Solution Let $\mathbf{x}, \mathbf{y} \in W$, $k \in \mathbb{R}$. Then there exist $c_i, d_i \in \mathbb{R}$ $(i=1, 2, \cdots, k)$ such that

$$\mathbf{x} = c_1\mathbf{x}_1+c_2\mathbf{x}_2+\cdots+c_k\mathbf{x}_k,\ \mathbf{y} = d_1\mathbf{x}_1+d_2\mathbf{x}_2+\cdots+d_k\mathbf{x}_k.$$

Hence

$$\mathbf{x}+\mathbf{y} = (c_1+d_1)\mathbf{x}_1+(c_2+d_2)\mathbf{x}_2+\cdots+(c_k+d_k)\mathbf{x}_k,$$

and $k\mathbf{x} = (kc_1)\mathbf{x}_1+(kc_2)\mathbf{x}_2+\cdots+(kc_k)\mathbf{x}_k$.

This implies $\mathbf{x}+\mathbf{y} \in W$, $k\mathbf{x} \in W$.

Hence, W is a subspace of $\mathbb{R}^n$. ■

In **Example 7**, we saw that for a subset $S=\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k\} \subseteq \mathbb{R}^n$, the set of all linear combinations $W=\{c_1\mathbf{x}_1+c_2\mathbf{x}_2+\cdots+c_k\mathbf{x}_k \mid c_1, c_2, \cdots, c_k \in \mathbb{R}\}$ of S is a subspace of $\mathbb{R}^n$. We say W is a **subspace of $\mathbb{R}^n$ spanned by S**. In this case, we say S **spans** W and S is a **spanning set** of W. We denote it

$$W=\text{span}(S) \quad \text{or} \quad W=\ <S>.$$

In particular, if all vectors in $\mathbb{R}^n$ can be expressed a linear combination of S, then S spans $\mathbb{R}^n$. That is,

$$\mathbb{R}^n =\ <S>\ = \{c_1\mathbf{x}_1+c_2\mathbf{x}_2+\cdots+c_k\mathbf{x}_k \mid c_1, c_2, \ldots, c_k \in \mathbb{R}\}$$

Example 8 (i) Show that $S=\{(1,1), (-1,2)\}$ is a spanning set of $\mathbb{R}^2$.

(ii) Show that $S=\{(1,0,0), (0,1,0), (1,-1,1)\}$ is a spanning set of $\mathbb{R}^3$.

Definition **[column space and row space]**

Let $A=[a_{ij}]\in M_{m\times n}$. Then, n columns $A^{(1)}, A^{(2)}, \cdots, A^{(n)}$ of A span a subspace of $\mathbb{R}^m$. This subspace is called a column space of A, denote by

$$< A^{(1)}, A^{(2)}, \ldots, A^{(n)} > \text{ or } \mathrm{Col}(A).$$

Similarly, a row space of A is defined by a subspace of $\mathbb{R}^n$ spanned by m rows $A_{(1)}, A_{(2)}, \cdots, A_{(m)}$ of A, denoted by

$$< A_{(1)}, A_{(2)}, \ldots, A_{(m)} > \text{ or } \mathrm{Row}(A).$$

Example 9

For

$$\mathbf{x}_1 = (1, 0, 1), \ \mathbf{x}_2 = (-3, 1, 1), \ \mathbf{x}_3 = (-2, 1, 2)$$

determine whether $S=\{\mathbf{x}_1, \ \mathbf{x}_2, \ \mathbf{x}_3\}$ spans $\mathbb{R}^3$ or not.

Solution This is a question whether there exist c_1, c_2, c_3 such that a given vector $\mathbf{x} = (x, y, z)$ is written as

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + c_3\mathbf{x}_3, \ (c_i \in \mathbb{R}).$$

(Using column vectors)

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = c_1\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} + c_3\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -3 & -2 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad \square$$

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(3, 3, [1, -3, -2, 0, 1, 1, 1, 1, 2])   # coefficient matrix
print A.rref()
```

```
[1 0 1]
[0 1 1]
[0 0 0]
```

This means that one of c_1, c_2, c_3 cannot be determined. Therefore this linear system has a case that the system cannot determine a unique solution. ■

Definition **[Linearly Independent and Linearly Dependent]**

If $S = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k\} \subseteq \mathbb{R}^n$ satisfies

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k = \mathbf{0} \quad (c_1, c_2, \ldots, c_k \in \mathbb{R})$$
$$\Rightarrow \; c_1 = c_2 = \cdots = c_k = 0$$

then $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$(or subset S) are called **linearly independent.**

If $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$(or subset S) are not linearly independent, then it is called **linearly dependent**.

- If S is linearly dependent, there exist at least one non-zero scalar in $\{c_1, c_2, \ldots, c_k\}$ such that

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k = \mathbf{0}.$$

The **unit vectors** of $\mathbb{R}^n$

$$\mathbf{e}_1 = (1, 0, \ldots, 0), \; \mathbf{e}_2 = (0, 1, \ldots, 0), \; \ldots, \; \mathbf{e}_n = (0, \ldots, 0, 1)$$

are linearly independent. This is because

$$c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_n\mathbf{e}_n = \mathbf{0}$$
$$\Rightarrow c_1(1, 0, \cdots, 0) + c_2(0, 1, \cdots, 0) + \cdots + c_n(0, 0, \cdots, 1) = (0, 0, \cdots, 0)$$
$$\Rightarrow (c_1, c_2, \cdots, c_n) = (0, 0, \cdots, 0)$$
$$\Rightarrow c_1 = c_2 = \cdots = c_k = 0.$$

Example 10 Show that for $\mathbf{x}_1 = (2, -1)$, $\mathbf{x}_2 = (1, 3)$, $S = \{\mathbf{x}_1, \mathbf{x}_2\}$ is linearly independent.

Solution For any $c_1, c_2 \in \mathbb{R}$,

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 = \mathbf{0} \;\Rightarrow\; c_1(2, -1) + c_2(1, 3) = (0, 0)$$
$$\Rightarrow\; 2c_1 + c_2 = 0, \; -c_1 + 3c_2 = 0$$

Thus $c_1 = c_2 = 0$, and S is linearly independent. ■

Example 11 Show that if $\mathbf{x}_1,\ \mathbf{x}_2,\ \mathbf{x}_3$ in $\mathbb{R}^n$ are linearly independent, then

$$\mathbf{x}_1,\ \mathbf{x}_1+\mathbf{x}_2,\ \mathbf{x}_1+\mathbf{x}_2+\mathbf{x}_3$$

are also linearly independent.

Solution For any $c_1, c_2, c_3 \in \mathbb{R}$,

$$c_1\mathbf{x}_1 + c_2(\mathbf{x}_1+\mathbf{x}_2) + c_3(\mathbf{x}_1+\mathbf{x}_2+\mathbf{x}_3) = \mathbf{0}$$
$$\Rightarrow\ (c_1+c_2+c_3)\mathbf{x}_1 + (c_2+c_3)\mathbf{x}_2 + c_3\mathbf{x}_3 = \mathbf{0}\ .$$

Since $\mathbf{x}_1,\ \mathbf{x}_2,\ \mathbf{x}_3$ are linearly independent,

$$c_1+c_2+c_3=0,\quad c_2+c_3=0,\quad c_3=0$$
$$\Rightarrow\ c_1=c_2=c_3=0$$

Therefore $\mathbf{x}_1,\ \mathbf{x}_1+\mathbf{x}_2,\ \mathbf{x}_1+\mathbf{x}_2+\mathbf{x}_3$ are linearly independent. ■

Example 12 For

$$\mathbf{x}_1 = (1, 0, -1),\ \mathbf{x}_2 = (1, 1, 0),\ \mathbf{x}_3 = (0, 1, 1)$$

in $\mathbb{R}^3$, Show that $S=\{\mathbf{x}_1,\ \mathbf{x}_2,\ \mathbf{x}_3\}$ is linearly dependent.

Solution For any $c_1, c_2, c_3 \in \mathbb{R}$, if $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + c_3\mathbf{x}_3 = \mathbf{0}$, then

$$c_1\begin{bmatrix}1\\0\\-1\end{bmatrix}+c_2\begin{bmatrix}1\\1\\0\end{bmatrix}+c_3\begin{bmatrix}0\\1\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}\quad\Rightarrow\quad\begin{bmatrix}1&1&0\\0&1&1\\-1&0&1\end{bmatrix}\begin{bmatrix}c_1\\c_2\\c_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$ □

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(3, 3, [1, 1, 0, 0, 1, 1, -1, 0, 1])  # coefficient matrix
print A.rref()
```

```
[ 1  0 -1]
[ 0  1  1]
[ 0  0  0]
```

This means that the above equations can be reduced to two equations of three variables. Since it has three variables more than the number of equations so that there are non-trivial solutions. One of them is given by $c_1 = 1$, $c_2 = -1$, $c_3 = 1$. Therefore there exist non zero scalars c_1, c_2, c_3, S is linearly dependent. ■

Theorem 3.4.1

For a set $S = \{\mathrm{x}_1, \mathrm{x}_2, \cdots, \mathrm{x}_k\} \subseteq \mathbb{R}^n$, the followings hold.

(1) A set S is linearly dependent if and only if some element in S can be expressed as a linear combination of the other elements in S.

(2) If S contains the zero vector, then S is a linearly dependent.

(3) If a subset S' of S is linearly dependent, then S is also linearly dependent.

If S is linearly independent, then S' is also linearly independent.

Proof (1) ($\Rightarrow$) If S is linearly dependent, then there exist $c_1, c_2, \cdots, c_k$ such that

$$c_1\mathrm{x}_1 + c_2\mathrm{x}_2 + \cdots + c_k\mathrm{x}_k = \mathbf{0}$$

where at least one element in $\{c_1, c_2, \cdots, c_k\}$ is a nonzero.
Without loss of generality, if $c_1 \neq 0$ then,

$$\mathrm{x}_1 = \left(-\frac{c_2}{c_1}\right)\mathrm{x}_2 + \cdots + \left(-\frac{c_k}{c_1}\right)\mathrm{x}_k$$

so that x_1 can be expressed as a linear combination of the other vectors in S

($\Leftarrow$) Without loss of generality, we can write

$$\mathrm{x}_1 = c_2\mathrm{x}_2 + \cdots + c_k\mathrm{x}_k$$

so that

$$(-1)\mathrm{x}_1 + c_2\mathrm{x}_2 + \cdots + c_k\mathrm{x}_k = \mathbf{0}$$

Hence, S is linearly dependent since $-1 \neq 0$.

Proofs of the rest are left as an exercise. ■

- In other words, that set S is linearly independent means that any vector in S cannot be written as a linear combination of the other vectors in S.

- In $\mathbb{R}^n$, there are at most n vectors in a linearly independent set.

Theorem 3.4.2 (For proof, see Theorem 7.1.2)

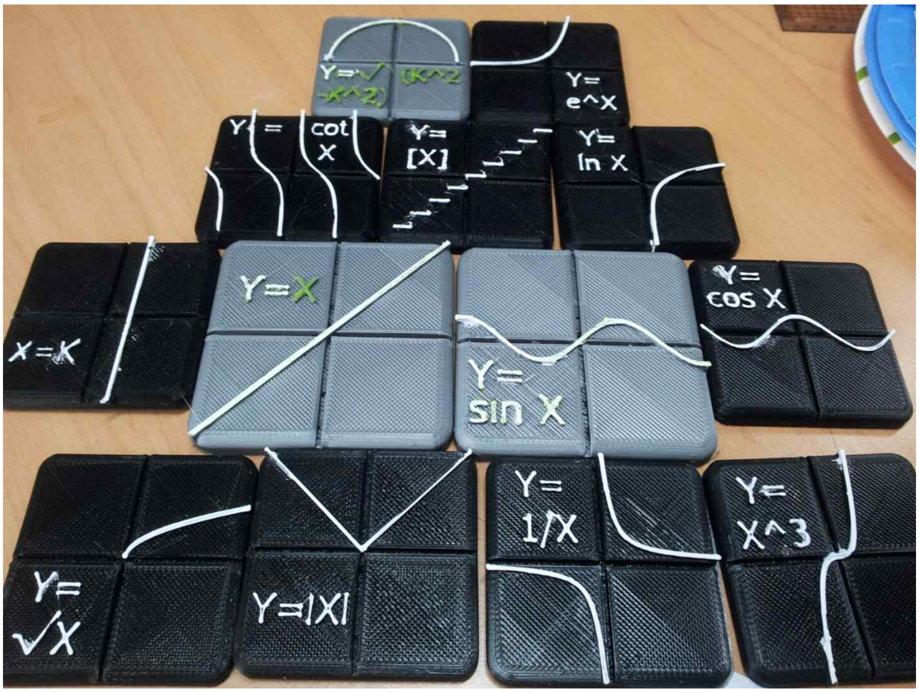
In $\mathbb{R}^n$, $m(>n)$ vectors are always linearly dependent.

Example 13 For $\mathbf{x}_1=(1,0,0)$, $\mathbf{x}_2=(1,1,0)$, $\mathbf{x}_3=(1,1,1)$, $\mathbf{x}_4=(0,1,1)$ in $\mathbb{R}^3$, we can easily check that $S=\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$ is linearly dependent from Theorem 3.4.2. ■

[Remark] Lines and plaines (from the viewpoint of subspace)

(1) Note that the span of nonzero vector $\mathbf{v}$ in $\mathbb{R}^n$. $\{t\mathbf{v} \mid t \in \mathbb{R}\}$ is a subspace containing the zero vector. Also $\{\mathbf{x}_0+t\mathbf{v} \mid t \in \mathbb{R}\}$ forms a line through $\mathbf{x}_0$ and parallel to $\mathbf{v}$. In other words, $\mathbf{x}=\mathbf{x}_0+t\mathbf{v}$ is translate of $\mathbf{x}=t\mathbf{v}$ by $\mathbf{x}_0$.

(2) In general, if $\mathbf{x}_0, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are vectors in $\mathbb{R}^n$, then $\mathbf{x}=\mathbf{x}_0+t_1\mathbf{v}_1+\cdots+t_k\mathbf{v}_k$ $(t_i \in \mathbb{R})$ is a subset of $\mathbb{R}^n$ which is the translation of a subspace $\mathbf{x}=t_1\mathbf{v}_1+\cdots+t_k\mathbf{v}_k$, through the origin, by $\mathbf{x}_0$.



3.5 Solution set and matrices

- Reference video: http://youtu.be/dalxHJBHL_g, http://youtu.be/O0TPCpKW_eY
- Practice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-4-Sec-3-5.html

In this section, we first state the relationship between invertibility of matrices and solutions to systems of linear equations, and then consider homogeneous systems.

Theorem 3.5.1 [Relation between an invertible matrix and its solution]

If an $n \times n$ matrix A is invertible and $\mathbf{b}$ is a vector in $\mathbb{R}^n$, the system

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

has a unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

Example 1 The following system can be written as $A\mathbf{x} = \mathbf{b}$.

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ 2x_1 + 5x_2 + 3x_3 &= 3 \\ x_1 \qquad + 8x_3 &= -1 \end{aligned}$$

where $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$. It is easy to show that A is invertible, and $A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$. Thus the solution of the above system is given by

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

That is $x_1 = -1$, $x_2 = 1$, $x_3 = 0$. □

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(3, 3, [1, 2, 3, 2, 5, 3, 1, 0, 8])        # coefficient matrix
b=vector([1, 3, -1])
Ai=A.inverse()                                     # inverse matrix calculation
print "x=", Ai*b
print
print "x=", A.solve_right(b)                       # solve directly.
```

x= (-1, 1, 0), x= (-1, 1, 0) ■

[Remark] **The homogeneous linear system**

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

can be written as $A\mathbf{x} = \mathbf{0}$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},\ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},\ \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The vector $\mathbf{x} = \mathbf{0}$ is called a **trivial solution**, and the solution $\mathbf{x} \neq \mathbf{0}$ is called a **nontrivial solution**. Since a homogeneous linear system always has a trivial solution, there are two cases as follows.

(1) It has only a trivial solution.
(2) It has infinitely many solutions (i.e. it has nontrivial solutions as well.)

Theorem 3.5.2 [Nontrivial solution of a homogeneous system]

A homogeneous system with m equations and n variables such that $m < n$ (i.e. the number of variables is greater than that of equations) has nontrivial solutions.

For a detailed proof for this theorem, see Linear Algebra : A Geometric Approach by S. Kumaresan, Prentice Hall of India, 2000.

Example 2 The homogeneous linear system

$$\begin{aligned} x_1 + \ x_2 + x_3 + \ x_4 &= 0 \\ x_1 + \qquad\qquad\quad x_4 &= 0 \\ x_1 + 2x_2 + x_3 \qquad &= 0 \end{aligned}$$

has the following augmented matrix and its RREF.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(3, 5, [1, 1, 1, 1, 0, 1, 0, 0, 1, 0, 1, 2, 1, 0, 0])
# augmented matrix
print "A="
print A
print
print "RREF(A)="
print A.rref()          # RREF
```

```
A=
[1 1 1 1 0]
[1 0 0 1 0]
[1 2 1 0 0]

RREF(A)=
[ 1  0  0  1  0]
[ 0  1  0 -1  0]
[ 0  0  1  1  0]
```

The corresponding system of equations is

$$\begin{aligned} x_1 + x_4 &= 0 \\ x_2 - x_4 &= 0 \\ x_3 + x_4 &= 0 \end{aligned}$$

Let $x_4 = r$ (r: a real number). Then the solution to (2) is

$$x_1 = r,\ x_2 = -r,\ x_3 = r,\ x_4 = r \quad (r \in \mathbb{R}).$$

The solution is trivial if $r = 0$, and nontrivial if $r \neq 0$. ■

Definition [The associated homogeneous system of linear equations]

Given a linear system $A\mathbf{x} = \mathbf{b}$, $A\mathbf{x} = \mathbf{0}$ is called the **associated homogeneous system of linear equations** of $A\mathbf{x} = \mathbf{b}$.

Example 3 Consider a system of linear equations.

$$\begin{bmatrix} 1 & 2 & -2 & 0 & 2 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 \\ 0 & 0 & 5 & 10 & 0 & 15 \\ 2 & 6 & 0 & 8 & 4 & 18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 5 \\ 6 \end{bmatrix}$$

The associated homogeneous linear system is as the following:

$$\begin{bmatrix} 1 & 2 & -2 & 0 & 2 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 \\ 0 & 0 & 5 & 10 & 0 & 15 \\ 2 & 6 & 0 & 8 & 4 & 18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Solution Since the matrix size is greater than 2, let us use Sage.

The RREF of the augmented matrix of the above system is as follows :

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(4, 7, [1, 2, -2, 0, 2, 0, 0, 2, 6, -5, -2, 4, -3, -1, 0, 0, 5, 10, 0,
15, 5, 2, 6, 0, 8, 4, 18, 6])          # augmented matrix
print A.rref()                          # RREF
```

```
[  1   0   0   4   2   0   0]
[  0   1   0   0   0   0   0]
[  0   0   1   2   0   0   0]
[  0   0   0   0   0   1 1/3]
```

Thus the above system reduces to

$$x_1 + 4x_4 + 2x_5 = 0,\ x_2 = 0,\ x_3 + 2x_4 = 0,\ x_6 = 1/3.$$

Note that x_4 and x_5 are free variables.

Let $x_4 = r,\ x_5 = s$. Then we have

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{3} \end{bmatrix} + r \begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix},\ r,\ s \in \mathbb{R}.$$

Consider the augmented matrix of RREF of its associated homogeneous linear system.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
B=matrix(4, 7, [1, 2, -2, 0, 2, 0, 0, 2, 6, -5, -2, 4, -3, 0, 0, 0, 5, 10, 0,
15, 0, 2, 6, 0, 8, 4, 18, 0])          # augmented matrix
print B.rref()                          # RREF
```

```
[1 0 0 4 2 0 0]
[0 1 0 0 0 0 0]
[0 0 1 2 0 0 0]
[0 0 0 0 0 1 0]
```

It is easy to see that the solution to this system is given by

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = r\begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix},\ r,\ s \in \mathbb{R}\,. \quad \blacksquare$$

When compared geometrically the solutions to a system and those of an associated homogeneous linear system, the solution set for the associated homogeneous linear system is merely translated by the vector $\mathbf{x}_0$ below.

$$\mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{3} \end{bmatrix}$$

We call the vector $\mathbf{x}_0$ a particular solution which can be obtained by substituting $r = s = 0$.

[Remark] **Relation between the solution set of the linear system and that of the associated homogeneous linear system.**

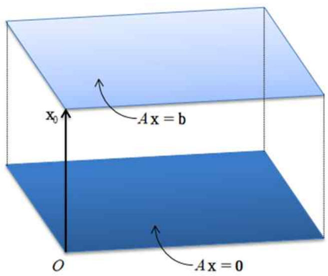
If $A\mathbf{x}=\mathbf{0}$ and $A\mathbf{x}_0=\mathbf{b}$, then

$$A(\mathbf{x}+\mathbf{x}_0)=A\mathbf{x}+A\mathbf{x}_0=\mathbf{0}+\mathbf{b}=\mathbf{b}.$$

Thus a system of linear equation $A\mathbf{x}=\mathbf{b}$ has solutions. Let W be a solution space to $A\mathbf{x}=\mathbf{0}$. If $\mathbf{x}_0$ is a solution to $A\mathbf{x}=\mathbf{b}$, then

$$\mathbf{x}_0+W=\{\mathbf{x}_0+\mathbf{x}\mid \mathbf{x}\in W\}$$

is a solution set of $A\mathbf{x}=\mathbf{b}$.



● A geometric meaning of $\mathbf{x}_0+W$ which is a solution set of $A\mathbf{x}=\mathbf{b}$ is a set of translation when a particular solution $\mathbf{x}_0$ is added to a solution set W of $A\mathbf{x}=\mathbf{0}$. Since $\mathbf{x}_0+W$ does not contain a zero vector, it is not a subspace of R^n.

Theorem 3.5.3 [Equivalent theorem of an invertible matrix]

For an $n\times n$ matrix A, the following are equivalent.

(1) $\mathrm{RREF}(A)=I_n$

(2) A is a product of elementary matrices.

(3) A is invertible.

(4) $\mathbf{0}$ is the unique solution to $A\mathbf{x}=\mathbf{0}$.

(5) $A\mathbf{x}=\mathbf{b}$ has a unique solution for any $\mathbf{b}\in\mathbb{R}^n$.

(6) The columns of A are linearly independent.

(7) The rows of A are linearly independent.

[Remark] **The vectors of the solution space of $A\mathbf{x}=\mathbf{0}$ are orthogonal to the rows of A.**

Let us think of the homogeneous system $A\mathbf{x}=\mathbf{0}$ with n variables. If the system has m linear equations, then the size of matrix A is $m\times n$. It can be rewritten using inner product. Let $A_{(1)}, A_{(2)}, \cdots, A_{(m)}$ indicate rows of a matrix A.

$$\begin{bmatrix} A_{(1)} \\ A_{(2)} \\ \vdots \\ A_{(m)} \end{bmatrix} \mathbf{x} = \begin{bmatrix} A_{(1)}\cdot \ \mathbf{x} \\ A_{(2)}\cdot \ \mathbf{x} \\ \vdots \\ A_{(m)}\cdot \ \mathbf{x} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Thus $A_{(i)}\cdot \ \mathbf{x}=0 (1\leq i\leq m)$ if $\mathbf{x}$ is a solution to $A\mathbf{x}=\mathbf{0}$. That is, the vectors in this solution space of $A\mathbf{x}=\mathbf{0}$ are all orthogonal to the row vectors of the matrix A.

Example 4 Consider the system of linear equations: $x_1+2x_2+x_3-3x_4=0$, $2x_1-x_2+x_3-2x_4=0$, $2x_1+x_2+x_3-3x_4=0$. It is easy to check that $\mathbf{v}=(1,1,3,2)$ is non-trivial solution of this system. Let us verify that $\mathbf{v}$ is orthogonal to row vectors of the coefficient matrix A of the above system.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix([[1,2,1,-3],[2,-1,1,-2],[2,1,1,-3]])
v=vector([1,1,3,2])
R=A.rows()
print v.dot_product(R[0])
print v.dot_product(R[1])
print v.dot_product(R[2])
```

```
0
0
0
```

Thus $\mathbf{v}$ is orthogonal to row vectors of the coefficient matrix A.

[Remark] **Hyperplane**

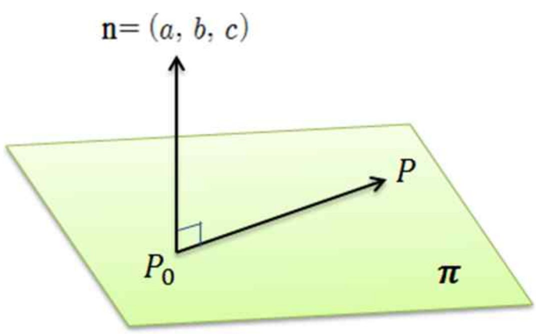
(1) Line of xy-plane: the solution set of a linear equation $a_1x+a_2y=b$, $(a_1, a_2) \neq (0,0)$

(2) Plane of xyz-space: the solution set of a linear equation $a_1x+a_2y+a_3z=b$, $(a_1, a_2, a_3) \neq (0,0,0)$

(3) Hyperplane of $\mathbb{R}^n$: the solution set of $a_1x_1+a_2x_2+\cdots+a_nx_n=b$, $\exists\ a_i \neq 0$
(If $b=0$, then it is a hyperplane passing through the origin)

$$\mathbf{a} \cdot \mathbf{x} = 0,\ (\mathbf{a} \neq \mathbf{0})$$

The set $\mathbf{a}^{\perp} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a} \cdot \mathbf{x} = 0\}$ is called an orthogonal complement of $\mathbf{a}$.



π: Orthogonal complement of n

3.6 Special matrices

- Reference video: http://youtu.be/dalxHJBHL_g, . http://youtu.be/jLh77sZOaM8
- Practice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-4-Sec-3-6.html

We saw various properties of matrix operations. In this section, we introduce special matrices and consider some of their crucial properties.

- Diagonal matrix: A square matrix with the entries 0 except the main diagonal. A diagonal matrix A with its main diagonal entries $a_{11}, a_{22}, \cdots, a_{nn}$ can be written as diag$(a_{11}, a_{22}, \ldots, a_{nn})$

$$\mathrm{diag}(a_{11}, a_{22}, \cdots, a_{nn}) = \begin{bmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix}$$

- Identity matrix: the matrix with its main diagonal entries all 1's, denoted by I_n

- Scalar matrix: kI_n

$$I_n = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}, \quad kI_n = \begin{bmatrix} k & & & \\ & k & & \\ & & \ddots & \\ & & & k \end{bmatrix}$$

Example 1 The following are all diagonal matrices. I and J are scalar matrices.

$$G = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}, \quad H = \begin{bmatrix} -3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad J = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

G and H are written as $G = \mathrm{diag}(2, -1)$ and $H = \mathrm{diag}(-3, -2, 1)$.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
G=diagonal_matrix([2, -1])           # generate diagonal matrix
H=diagonal_matrix([-3, -2, 1])       # diagonal_matrix([a1, a2, a3])
print G
print H
```

```
[ 2  0]          [-3  0  0]
[ 0 -1]          [ 0 -2  0]
                 [ 0  0  1]
```
■

Example 2 Consider the following matrix.

If $D=\begin{bmatrix}1&0&0\\0&-3&0\\0&0&2\end{bmatrix}$ and $A=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}$, $DA=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\-3a_{21}&-3a_{22}&-3a_{23}\\2a_{31}&2a_{32}&2a_{33}\end{bmatrix}$.

For a general matrix $A=[a_{ij}]_{n\times n}$, DA is obtained by multiplying each row of A by the corresponding entry of D, and AD is obtained by multiplying each column of A by the corresponding entry of D,

Furthermore, it satisfies the following.

$$D^{-1}=\begin{bmatrix}1&0&0\\0&-\frac{1}{3}&0\\0&0&\frac{1}{2}\end{bmatrix},\ D^5=\begin{bmatrix}1&0&0\\0&-243&0\\0&0&32\end{bmatrix},\ D^{-5}=\begin{bmatrix}1&0&0\\0&-\frac{1}{243}&0\\0&0&\frac{1}{32}\end{bmatrix}$$

In other words, the power of a diagonal matrix is the same as the diagonal matrix with the powers of the entries of the main diagonal. □

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
D=diagonal_matrix([1, -3, 2])          # generating a diagonal matrix
print "D^(-1)="
print D^(-1)
print
print "D^5="
print D^5
```

```
D^(-1)=
[   1    0    0]
[   0 -1/3    0]
[   0    0  1/2]

D^5=
[   1    0    0]
[   0 -243    0]
[   0    0   32]
```
■

Definition

If a square matrix A satisfies $A^T = A$, A is called a symmetric matrix.
If $A^T = -A$, then A is called a skew-symmetric matrix.

Example 3 In the following matrices, A and I_3 are symmetric matrices and B is a skew-symmetric matrix.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- http://matrix.skku.ac.kr/RPG_English/3-SO-Symmetric-M.html

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(3, 3, [1, 2, 3, 2, 4, 5, 3, 5, 6])
B=matrix(3, 3, [0, 1, -2, -1, 0, 3, 2, -3, 0])
print bool(A==A.transpose())          # Check if A symmetric
print bool(-B==B.transpose())         # Check if B anti-symmetric
```

```
True
True
```

■

Example 4 If A is a square matrix, prove the following.

(1) $A + A^T$ is a symmetric matrix.

(2) $A - A^T$ is a skew-symmetric matrix.

Solution (1) Since $(A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T$, $A + A^T$ is a symmetric matrix.

(2) Since $(A - A^T)^T = A^T - (A^T)^T = A^T - A = -(A - A^T)$, $A - A^T$ is a skew-symmetric matrix. ■

[Remark]

A given matrix can be written uniquely as a sum of a symmetric matrix and a skew-symmetric matrix.

Proof For any given matrix A, $A = \dfrac{A+A^T}{2} + \dfrac{A-A^T}{2}$ and $\dfrac{A+A^T}{2}$ is a symmetric matrix and $\dfrac{A-A^T}{2}$ is a skew-symmetric matrix. ■

- Lower triangular matrix: A square matrix whose entries under the main diagonal are all zeros
- Upper triangular matrix: A square matrix whose entries above the main diagonal are all zeros

Example 5 In general, 4×4 triangular matrices are as follows.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix} : \text{Upper Triangular Matrix} \qquad \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} : \text{Lower Triangular Matrix}$$

Theorem 3.6.1 [Property of a triangular matrix]

Let A and B be a lower triangular matrix.

(1) $A\cdot B$ is a lower triangular matrix.

(2) If A is an invertible matrix, then A^{-1} is a lower triangular matrix.

(3) If $a_{ii}=1$ for all i, then the main diagonal entries of A^{-1} is all 1's.

Example 6 Let A be a square matrix. If there exists an positive integer k such that $A^k = O$ (A is called nilpotent), $(I-A)$ is invertible and $(I-A)^{-1} = I + A + A^2 + \cdots + A^{k-1}$. This is because

$$(I-A)(I+A+\cdots+A^{k-1}) = I$$ ■

Chapter 3 Exercises

- http://matrix.skku.ac.kr/LA-Lab/index.htm
- http://matrix.skku.ac.kr/knou-knowls/cla-sage-reference.htm

T/F Problem Indicate whether the statement is true (T) or false (F). Justify your answer.

(a) If three nonzero vectors form a linearly independent set, then each vector in the set can be expressed as a linear combination of the other two.

(b) The set of all linear combinations of two vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^n$ is a plane.

(c) If $\mathbf{u}$ cannot be expressed as a linear combination of $\mathbf{v}$ and $\mathbf{w}$, then the three vectors are linearly independent.

(d) A set of vectors in $\mathbb{R}^n$ that contains is linearly dependent.

(e) If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly independent set, then so is the set $\{k\mathbf{v}_1, k\mathbf{v}_2, k\mathbf{v}_3\}$ for every nonzero scalar k.

Problem 1 When $A = \begin{bmatrix} 1 & 3 \\ 4 & -1 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 2 & 5 \\ 1 & -1 & 4 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 2 \end{bmatrix}$, confirm the following.

$A(BC) = (AB)C$.

Problem 2 When $A = \begin{bmatrix} -2 & 3 \\ 2 & -3 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 3 \\ 2 & 0 \end{bmatrix}$, $C = \begin{bmatrix} -4 & -3 \\ 0 & -4 \end{bmatrix}$, confirm that $AB = AC$ and that $B \neq C$.

Problem 3 When $A = \begin{bmatrix} 1 & 3 \\ 4 & -1 \end{bmatrix}$, compute the following.

$3A^3 - 2A^2 + 5A - 4I_2$

Problem 4 Show that B is the inverse of A. And confirm that $(A^T)^{-1} = (A^{-1})^T$.

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & -3 \\ 3 & 6 & -5 \end{bmatrix},\ B = \begin{bmatrix} 2 & -17 & 11 \\ -1 & 11 & -7 \\ 0 & 3 & -2 \end{bmatrix}$$

Problem 5 If $A^2 = A$, show that $(I-2A) = (I-2A)^{-1}$.

Problem 6 Find a 3×3 elementary matrix corresponding to each elementary operation.

(1) $R_2 \leftrightarrow R_3$

(2) $2R_2 \rightarrow R_2$

(3) $-2R_1 + R_3 \rightarrow R_3$

Problem 7 Using elementary operations, find the inverse of the following matrix.

(1) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ (2) $\begin{bmatrix} 1 & 4 & 3 & 2 \\ 7 & 26 & 20 & 13 \\ 1 & 2 & 3 & 4 \\ 5 & 4 & 6 & 3 \end{bmatrix}$

Solution $[A \vdots I_4] = \left[\begin{array}{cccc|cccc} 1 & 4 & 3 & 2 & 1 & 0 & 0 & 0 \\ 7 & 26 & 20 & 13 & 0 & 1 & 0 & 0 \\ 1 & 2 & 3 & 4 & 0 & 0 & 1 & 0 \\ 5 & 4 & 6 & 3 & 0 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cccc|cccc} 1 & 4 & 3 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} & \frac{7}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 3 & 6 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 & \frac{57}{4} & -\frac{9}{4} & \frac{1}{4} & \frac{1}{4} \end{array}\right] \rightarrow$

$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & \frac{95}{4} & -\frac{15}{4} & -\frac{1}{4} & \frac{3}{4} \\ 0 & 1 & 0 & 0 & \frac{59}{4} & -\frac{9}{4} & -\frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 1 & 0 & -\frac{147}{4} & \frac{23}{4} & \frac{1}{4} & -\frac{3}{4} \\ 0 & 0 & 0 & 1 & \frac{57}{4} & -\frac{9}{4} & \frac{1}{4} & \frac{1}{4} \end{array}\right] = [C \vdots D]$. $D = A^{-1} = \begin{bmatrix} \frac{95}{4} & -\frac{15}{4} & -\frac{1}{4} & \frac{3}{4} \\ \frac{59}{4} & -\frac{9}{4} & -\frac{1}{4} & \frac{1}{4} \\ -\frac{147}{4} & \frac{23}{4} & \frac{1}{4} & -\frac{3}{4} \\ \frac{57}{4} & -\frac{9}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$. ■

Problem 8 Let $E=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$ and A be any 3×3 matrix.

(1) What is EA and confirm how E affects on A.

(2) What is AE and confirm how E affects on A.

Problem 9 Determine if W is a subspace of $\mathbb{R}^2$.

$W_2 = \{(x_1, x_2) \mid x_1 x_2 = 0\}$

Problem 10 Determine if W is a subspace of $\mathbb{R}^3$.

$W_6 = \{(x_1, x_2, x_3) \mid x_1 = x_2 = x_3\}$

Problem 11 Find a vector equation and a parameterized equation of the subspace spanned by the following vectors.

(a) $\mathrm{v}_1 = (4, -4, 2)$, $\mathrm{v}_2 = (-3, 5, 7)$

(b) $\mathrm{v}_1 = (1, 5, -1, 4, 2)$, $\mathrm{v}_2 = (2, 2, 0, 1, -4)$

Solution (a) $x_1 = 4s - 3t$, $x_2 = -4s + 5t$, $x_3 = 2s + 7t$ where s, t in $\mathbb{R}$.

(b) $x_1 = s + 2t$, $x_2 = 5s + 2t$, $x_3 = -s$, $x_4 = 4s + t$, $x_5 = 2s - 4t$. ■

Problem 12 Give a solution by finding the inverse of the coefficient matrix of the system.

$$\begin{cases} 3x \qquad - z = 1 \\ 3x + 4y - 2z = 1 \\ 3x + 5y - 2z = 2 \end{cases}$$

Problem 13 Determine if the homogeneous system has a nontrivial solutoin.

$$\begin{cases} x+y+\ z-\ w=0 \\ x-y+2z+\ w=0 \\ x\qquad -\ z-5w=0 \end{cases}$$

Problem 14 Check if the following matrix is invertible. If so, find its inverse by using a property of special matrices.

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Problem 15 Find the product by using a property of special matrices.

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

Problem 16 Determine a, b, c, d so that A is skew-symmetric matrix.

$$A = \begin{bmatrix} 0 & 2a & 3a-3b \\ -2 & 0 & 2a-4c \\ -6 & -5 & d \end{bmatrix}$$

PI If $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ satisfies $a_{11} \neq 0$ and $\alpha = \dfrac{a_{21}}{a_{11}}$, show that A can be expressed as follows.

$$\begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ 0 & b \end{bmatrix}$$

What is the value of b?

Solution $A = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ 0 & b \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ => $a_{22} = \alpha a_{11} + a_{12} b$

$$\therefore \quad b = a_{22} - \frac{a_{12}a_{21}}{a_{11}} = \frac{a_{11}a_{22} - a_{12}a_{21}}{a_{11}}$$ ■

P2 Let A be a square matrix. Explain why the following hold.

(1) If A contains a row or a column consisting of 0's, A is not invertible.

(2) If A contains the same rows or columns, A is not invertible.

(3) If A contains a row or column which is a scalar multiple of another row or column of A.

P3 Let A be an $n\times n$ square matrix. Discuss what condition is need to have $AB=AC \Rightarrow B=C$.

P4 Find 2×2 matrices A, B and explain the relation with ERO.

$A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $B=\begin{bmatrix} d & c \\ b & a \end{bmatrix}$

P5 Decide if the following 4 vectors are linearly independent.

$\mathrm{v}_1=(4,\ -5,\ 2,\ 6)$, $\mathrm{v}_2=(2,\ -2,\ 1,\ 3)$, $\mathrm{v}_3=(6,\ -3,\ 3,\ 9)$, $\mathrm{v}_4=(4,\ -1,\ 5,\ 6)$

P6 If $A\mathrm{x}=\mathrm{b}$ and $A\mathrm{x}=\mathrm{c}$ have a solution, prove that $A\mathrm{x}=\mathrm{b}+\mathrm{c}$ has a solution.

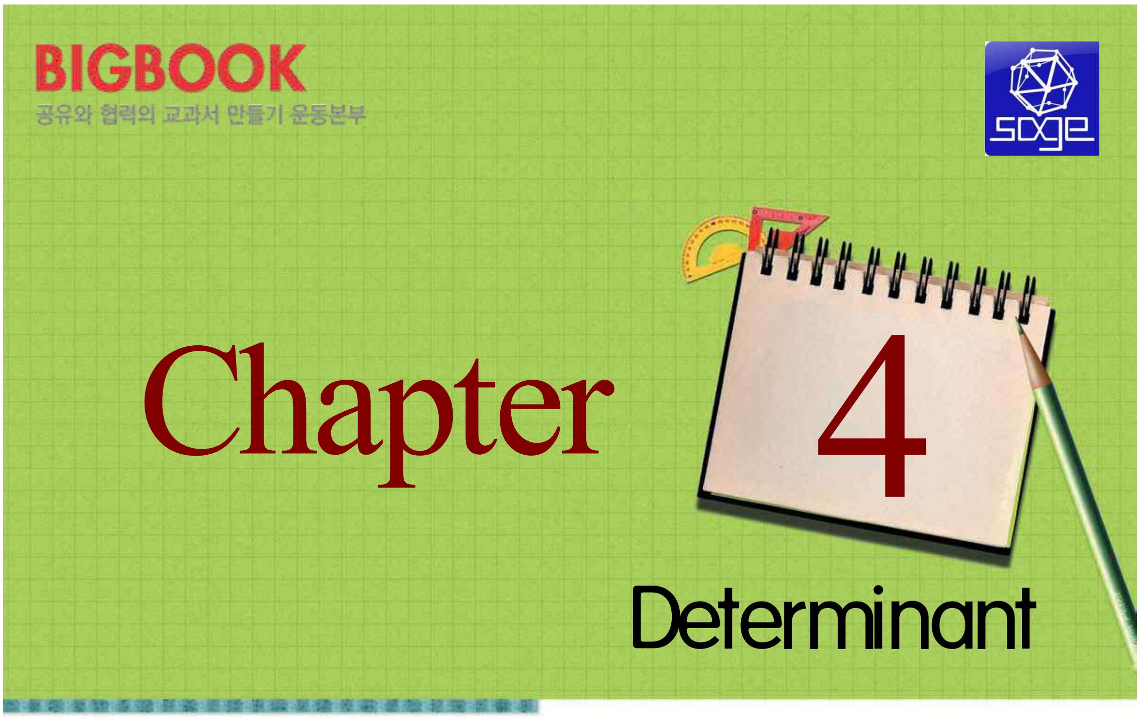
P7 Suppose A is an invertible matrix of order n. If v in $\mathbb{R}^n$ is perpendicular to every row of A, what is v? Justify your answer.

P8 Prove that a necessary and sufficient condition for a diagonal matrix to be invertible is that there is no zero entry in the main diagonal.

P9 If A is invertible and symmetric, so is A^{-1}.

Solution $A=A^T$, $AA^{-1}=I$ and $I=I^T=(AA^{-1})^T=(A^{-1})^TA^T$.

=> $(A^{-1})^TA=I$ => $A^{-1}=(A^{-1})^T$ => A^{-1} is symmetric. ■


BIGBOOK
공유와 협력의 교과서 만들기 운동본부
Chapter
4
Determinant

The concept of **determinant** was introduced 150 years before the use of modern matrix, and we have used the determinant to solve the systems of linear equations for over 100 years. In late 19th century, Sylvester introduced **the concept of matrix** and **the method for solving systems of equations by using an inverse matrix, where the determinant is used to check if an inverse of a matrix exists or not.** Also, the determinant can be used to find area, volume, equations of lines or planes, and exterior product. It also helps in **geometric interpretation of vectors**.

In this chapter, we first define the determinant and review its properties. Then we study how to compute the determinant by cofactor expansion. We also study Cramer's rule which solves the systems of linear equations by using the determinant.

One of the most important concepts in linear algebra is **eigenvalues** and **eigenvectors**. Eigenvalues have almost all important informations by n components from an object with n^2 components. Eigenvalues are not only important in theoretical perspective but also applicable to almost all areas related to matrix, such as, finding the solutions of differential equations, computing the power of given matrix, Google search, and image compression, etc. In the last section of this chapter, we compute eigenvalues by using the determinant.

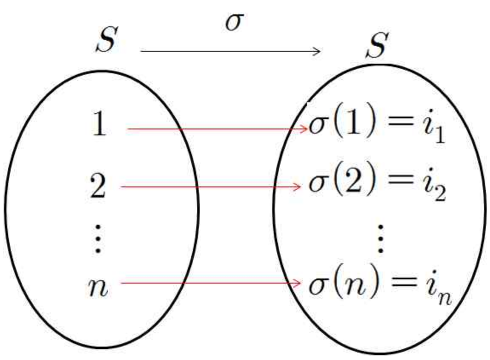
4.1 Definition of Determinant

- Reference video: http://youtu.be/DM-q2ZuQtI0, http://youtu.be/Vf8LlkKKHgg
- Practice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-5-Sec-4-1.html

In this section, we introduce a determinant function which assigns any square matrix A to a real number $f(A)$. In order to define the determinant function, we first introduce permutation. Then we review the properties of the determinant function.

Definition [Permutation]

For a set of natural numbers $S=\{1, 2, \ldots, n\}$, **permutation** is a one to one function from S to S.



- We simply denote a permutation as $\sigma=(\sigma(1)\ \sigma(2)\ \cdots\ \sigma(n))=(i_1\ \ i_2\ \cdots\ i_n)$. As a permutation σ is an one to one correspondence, the range $\{i_1, i_2, \ldots, i_n\}$ is simply a rearrangement of $1, 2, \ldots, n$. Hence, there are $n!$ permutations on $S=\{1, 2, \cdots, n\}$. We denote the set of all permutations of set S by S_n.

n	S_n	$n!$
1	(1)	$1!$
2	$(1\ 2)$, $(2\ 1)$	$2!$
3	$(1\ 2\ 3)$, $(2\ 3\ 1)$, $(3\ 1\ 2)$, $(1\ 3\ 2)$, $(2\ 1\ 3)$, $(3\ 2\ 1)$	$3!$
4	$(1\ 2\ 3\ 4)$, $(1,2,4,3)$, ... , $(4,3,2,1)$, $(4\ 3\ 2\ 1)$	$4!$

[Remark] Inversion

In permutation $(j_1 \ \ j_2 \ \cdots \ j_n)$, an **inversion** is the case when a bigger natural number placed **on the left hand side** of a smaller natural number. For example, in a permutation $(1 \ \ 4 \ \ 2 \ \ 3)$, 4 is placed on the left hand side of 2, and hence $(4 \ \ 2)$ is an inversion. Similarly, $(4 \ \ 3)$ is an inversion.

$$(1 \ \ 4 \ \ 2 \ \ 3)$$

- **Number of inversions for j_k** : after $(k+1)$-th index, the number of indexes which is smaller than k-th index j_k is called **the number of inversions for** j_k. In the above example, the number of inversion for 4 is 2. **Number of inversions for a permutation $(j_1 \ \ j_2 \ \cdots \ j_n)$ is the total sum of each number of inversions for j_k,** $k = 1, 2, \cdots, n$.

Definition **[Even permutation and odd permutation]**

If number of inversions for a permutation is even than it is called an **even permutation**, If the number is odd than it is called an **odd permutation**.

Example 1

Determine whether it is even or odd permutation by computing the inversion numbers for a permutation $\sigma = (5 \ \ 1 \ \ 2 \ \ 4 \ \ 3)$ in S_5.

Solution The number of inversions for 5 is 4. The number of inversions for 1 is 0, for 2 is 0, for 4 is 1, and 3 is the last index. Hence, the total sum is $4 + 0 + 0 + 1 + 0 = 5$, and it is an odd permutation. □

- http://matrix.skku.ac.kr/RPG_English/4-TF-Permutation.html

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
Permutation([5,1,2,4,3]).inversions()          # inversions
```

[[0, 1], [0, 2], [0, 3], [0, 4], [3, 4]]   # Note!! Index starts from 0

Permutation([5,1,2,4,3]).number_of_inversions()   # Number of inversions

5

Permutation([5,1,2,4,3]).is_even() # check whether it is even permutation

False ■

Definition **[Signature function]**

Signature function $\mathrm{sgn}: S_n \to \{+1, -1\}$, which assigns each permutation of S_n to either +1 or -1 as follows.

$$\mathrm{sgn}(\sigma) = \begin{cases} +1 & (\sigma : \text{even permutation}) \\ -1 & (\sigma : \text{odd permutation}) \end{cases}$$

Example 2 Classify the permutations of S_3 to either even or odd permutation.

Solution

permutations	number of inversions	class	sign
$(1\ 2\ 3)$	0	even	$+$
$(2\ 3\ 1)$	$2 \quad (2\ 1, 3\ 1)$	even	$+$
$(3\ 1\ 2)$	$2 \quad (3\ 1, 3\ 2)$	even	$+$
$(1\ 3\ 2)$	$1 \quad (3\ 2)$	odd	$-$
$(2\ 1\ 3)$	$1 \quad (2\ 1)$	odd	$-$
$(3\ 2\ 1)$	$3 \quad (3\ 2,\ 3\ 1,\ 2\ 1)$	odd	$-$

■

- In permutation, if two numbers switch the location then the signature is changed

Theorem 4.1.1

Let τ be a permutation by switching any two numbers from given permutation σ. Then

$$\text{sgn}(\tau) = -\text{sgn}(\sigma)$$

Definition [Determinant] [Leibniz formula]

Let $A = [a_{ij}]$ be an $n \times n$ matrix. We denote the determinant of matrix A as $\det(A)$ or $|A|$ and define it as follows.

$$\det(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

- By definition, 1×1 matrix $A = [a]$ has it's determinant as $\det(A) = a$.

- Each term $\text{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$ in the determinant is from the matrix A, by choosing a row and a column, without any overlapping, then multiplying them and assigning a corresponding signature.

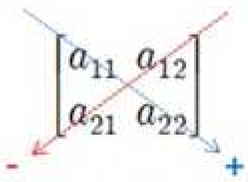
Example 3

Find the $\det(A)$, where $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$.

Solution

As A is 2×2 matrix, $S_2 = \{\sigma_1, \sigma_2\} = \{(1\ 2), (2\ 1)\}$. Since $\text{sgn}(\sigma_1) = 1$, $\text{sgn}(\sigma_2) = -1$, we have

$\det(A) = \text{sgn}(\sigma_1) a_{1\sigma_1(1)} a_{2\sigma_1(2)} + \text{sgn}(\sigma_2) a_{1\sigma_2(1)} a_{2\sigma_2(2)} = a_{11}a_{22} - a_{12}a_{21}$.



Sarrus Method

https://en.wikipedia.org/wiki/Rule_of_Sarrus ■

Example 4

Find the $\det(A)$, where $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$.

Solution

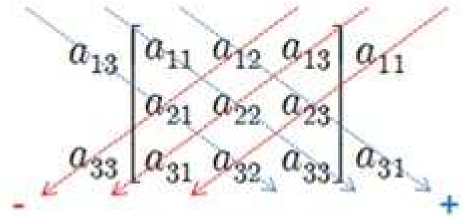
As A is 3×3 matrix,

$$S_3 = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6\} = \{(1\ 2\ 3),\ (2\ 3\ 1),\ (3\ 1\ 2),\ (1\ 3\ 2),\ (2\ 1\ 3),\ (3\ 2\ 1)\}.$$

Since $\operatorname{sgn}(\sigma_1) = 1$, $\operatorname{sgn}(\sigma_2) = 1$, $\operatorname{sgn}(\sigma_3) = 1$, $\operatorname{sgn}(\sigma_4) = -1$, $\operatorname{sgn}(\sigma_5) = -1$, $\operatorname{sgn}(\sigma_6) = -1$,

by substituting them into the definition of the determinant, we have

$$\begin{aligned} \det(A) &= \operatorname{sgn}(\sigma_1) a_{1\sigma_1(1)} a_{2\sigma_1(2)} a_{3\sigma_1(3)} + \operatorname{sgn}(\sigma_2) a_{1\sigma_2(1)} a_{2\sigma_2(2)} a_{3\sigma_2(3)} \\ &\quad + \cdots + \operatorname{sgn}(\sigma_6) a_{1\sigma_6(1)} a_{2\sigma_6(2)} a_{3\sigma_6(3)} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \end{aligned}$$



Sarrus Method ■

Example 5

Compute the determinant of the following matrices.

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

Solution

$$|A| = \begin{vmatrix} 3 & 1 \\ 4 & -2 \end{vmatrix} = 3(-2) - (1)(4) = -10.$$

$$|B| = \begin{array}{c|ccc|c} 3 & 1 & 2 & 3 & 1 \\ & -4 & 5 & 6 & \\ 9 & 7 & -8 & 9 & 7 \end{array} = (45) + (84) + (96) - (105) - (-48) - (-72) = 240. \quad \square$$

- http://matrix.skku.ac.kr/RPG_English/4-B1-Det-matrix.html

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
B=matrix(QQ, 3, 3, [1, 2, 3, -4, 5, 6, 7, -8, 9])
print B.det()                          # compute the determinant
```

```
240
```

■

[Remark] Sarrus' method cannot be applied to the case of degree 4 or higher.

Hence, the determinant with degree 4 or higher should be computed by the definition. But in that case, there are too many terms and signs to be determined. (Indeed, for degree 4 case, there are $4! = 24$ terms, and for degree 10, there are $10! = 3{,}628{,}800$ terms to compute). Therefore, it is more effective to study the properties of the determinant and find the determinant by using these properties. (We will skip the proofs but will verify them by examples).

Properties of the determinant

Theorem 4.1.2

A square matrix A and its transpose matrix A^T have the same determinant.

http://math.stackexchange.com/questions/123923/a-matrix-and-its-transpose-have-the-same-set-of-eigenvalues

Example 6 In Example 5, $|B| = 240$, and $B^T = \begin{bmatrix} 1 & -4 & 7 \\ 2 & 5 & -8 \\ 3 & 6 & 9 \end{bmatrix}$. Since

$$|B^T| = \begin{vmatrix} 1 & -4 & 7 \\ 2 & 5 & -8 \\ 3 & 6 & 9 \end{vmatrix} = (45) + (96) + (84) - (105) - (-48) - (-72) = 240$$

we have $|B| = |B^T|$. □

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
B=matrix(QQ, 3, 3, [1, 2, 3, -4, 5, 6, 7, -8, 9])
print B.transpose().det()
```

240 ■

- The properties of the determinant regarding to rows also work to columns.

Theorem 4.1.3

Let B be a matrix obtained by switching two rows (columns) from a square matrix A then $|B| = -|A|$.

Proof Let $B = [b_{ij}]$ be a matrix obtained by replacing r th and s th row of $A = [a_{ij}]$, $r < s$. This means $b_{rj} = a_{sj}$, $b_{sj} = a_{rj}$ and $b_{ij} = a_{ij}$ if $i \neq r, s$.

$$
\begin{aligned}
|B| &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, b_{1\sigma(1)} \cdots b_{r\sigma(r)} \cdots b_{s\sigma(s)} \cdots b_{n\sigma(n)} \quad \text{(by definition)} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{s\sigma(r)} \cdots a_{r\sigma(s)} \cdots a_{n\sigma(n)} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{r\sigma(s)} \cdots a_{s\sigma(r)} \cdots a_{n\sigma(n)} \\
&= -\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{r\sigma(r)} \cdots a_{s\sigma(s)} \cdots a_{n\sigma(n)} \quad \text{(by theorem 4.1.1)} \\
&= -|A|
\end{aligned}
$$
■

Example 7 Let $A = \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 2 \\ 2 & -1 \end{bmatrix}$. Since $\begin{vmatrix} 2 & -1 \\ 3 & 2 \end{vmatrix} = 7$ and $\begin{vmatrix} 3 & 2 \\ 2 & -1 \end{vmatrix} = -7$, $|B| = -|A|$. ■

Theorem 4.1.4

If a square matrix A has two identical rows (columns) then $|A| = 0$.

Example 8 Let $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 7 \\ 1 & 2 & 3 \end{bmatrix}$ which has identical first and third rows. Note

$$|A| = \begin{vmatrix} 1 & 2 & 3 \\ -1 & 0 & 7 \\ 1 & 2 & 3 \end{vmatrix} = (0)+(14)+(-6)-(0)-(14)-(-6) = 0$$

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ, 3, 3, [1, 2, 3, -1, 0, 7, 1, 2, 3])
print A.det()                              # compute the determinant
```

```
0
```

■

Theorem 4.1.5

If a square matrix A has a row (column) with identical zeros then $|A| = 0$.

Example 9

Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 0 \end{bmatrix}$ which has identical zeros in the third row. Observe

$$|A| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 0 \end{vmatrix}$$

$$= 1\times 5\times 0 + 2\times 6\times 0 + 3\times 4\times 0 - 2\times 4\times 0 - 3\times 5\times 0 - 1\times 6\times 0 = 0$$ ■

Theorem 4.1.6

Let B be a matrix obtained by multiplying k times a row of a square matrix A. Then $|B| = k|A|$.

Example 10

Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 5 & 3 \\ 2 & 8 & 6 \end{bmatrix}$. Note that

$$|A| = \begin{vmatrix} 1 & 2 & 3 \\ 1 & 5 & 3 \\ 2 & 8 & 6 \end{vmatrix} = 2\begin{vmatrix} 1 & 2 & 3 \\ 1 & 5 & 3 \\ 1 & 4 & 3 \end{vmatrix} = (2)(3)\begin{vmatrix} 1 & 2 & 1 \\ 1 & 5 & 1 \\ 1 & 4 & 1 \end{vmatrix} = (2)(3)(0) = 0$$ ■

Theorem 4.1.7

If a square matrix A has two proportional rows then $|A| = 0$.

Theorem 4.1.8

Let A be a square matrix and k times of one row is added to another row of A and name this new matrix as B, then $|B| = |A|$.

Proof Let B be a new matrix whose second row is obtained by adding k times of the first row of A to $A \in M_n$.

$$\det(B) = \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} (k a_{1\sigma(2)} + a_{2\sigma(2)}) a_{3\sigma(3)} \cdots a_{n\sigma(n)}$$

$$\det(B) = k \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{1\sigma(2)} a_{3\sigma(3)} \cdots a_{n\sigma(n)} + \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

$$\Rightarrow \quad \det(B) = \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} = |A| \quad \text{(by Theorem 4.1.4)}$$ ■

Example 11 Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 3 \\ 1 & 0 & 1 \end{bmatrix}$ and 2 times of the second row is added to the first row and name it as matrix B. Then

$$B = \begin{bmatrix} 5 & 0 & 9 \\ 2 & -1 & 3 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1+2\cdot 2 & 2+2(-1) & 3+2\cdot 3 \\ 2 & -1 & 3 \\ 1 & 0 & 1 \end{bmatrix}.$$

Note that $|A| = 4 = |B|$. ■

Theorem 4.1.9

If $A = [a_{ij}]$ is an $n \times n$ triangular matrix, the determinant of A equals the product of the diagonal elements. That is,

$$|A| = a_{11} a_{22} \cdots a_{nn}$$

Example 12 From the previous theorem, $\begin{vmatrix} 2 & 7 & -3 & 8 & 3 \\ 0 & -3 & 7 & 5 & 1 \\ 0 & 0 & 6 & 7 & 6 \\ 0 & 0 & 0 & 9 & 8 \\ 0 & 0 & 0 & 0 & 4 \end{vmatrix} = (2)(-3)(6)(9)(4) = -1296$. ■

[Remark] **How to compute the determinant**

1. Use elementary row operations to make many zeros to a certain row (column).
2. Multiply the diagonal elements.

※ Note that during the elementary row operations, if you multiply k times a row (column) or switch two rows (columns), do not forget to multiply $1/k$ and -1.

Example 13 Find the determinant of a matrix A, where

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

Solution

$$|A| = \begin{vmatrix} 0 & 1 & 5 \\ 3 & -6 & 9 \\ 2 & 6 & 1 \end{vmatrix} = -\begin{vmatrix} 3 & -6 & 9 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{vmatrix} \qquad R_1 \leftrightarrow R_2$$

$$= -3\begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{vmatrix}$$

$$= -3\begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 10 & -5 \end{vmatrix} \qquad (-2)R_1 + R_3$$

$$= -3\begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & -55 \end{vmatrix} \qquad (-10)R_2 + R_3$$

$$= (-3)(-55)\begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{vmatrix} = (-3)(-55)(1) = 165$$

■

Theorem 4.1.10

Let E be an $n \times n$ elementary matrix. Then $\det(EA) = \det(E)\det(A)$.

[Remark] **The determinant of an elementary matrix**

1. If E is obtained by multiplying k $(k \neq 0)$ to a row of I_n, $\det(E) = k$
2. If E is obtained by switching two rows of I_n, $\det(E) = -1$
3. If E is obtained by multiplying k times a row and adding it to another row of I_n, $\det(E) = 1$
4. If A is an $n \times n$ matrix and E is an elementary matrix, $\det(EA) = \det(E) \cdot \det(A)$.

● **Equivalent conditions for invertible matrix**

Theorem 4.1.11

A is invertible if and only if $\det A \neq 0$.

Theorem 4.1.12

For any two $n \times n$ matrices A and B, $|AB| = |A||B|$.

Example 14 Verify the above theorem with matrices $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$.

Solution Since $AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 10 & 5 \end{bmatrix}$, and

$|A| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = -2$, $|B| = \begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} = 5$, $|AB| = \begin{vmatrix} 4 & 3 \\ 10 & 5 \end{vmatrix} = -10$

$|AB| = -10 = |A||B|$. □

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ, 2, 2, [1, 2, 3, 4])
B=matrix(QQ, 2, 2, [2, -1, 1, 2])
C=A*B
print "det(AB)=", C.det()
print "det(A)*det(B)=", A.det()*B.det()
```

```
det(AB)= -10
det(A)*det(B)= -10
```
■

Theorem 4.1.13

If a square matrix A is invertible then $|A| \neq 0$ and $|A^{-1}| = \dfrac{1}{|A|}$.

Example 15

Verify the above Theorem with a matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

Solution A is invertible with $A^{-1} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$. Observe $|A| = -2 \neq 0$ and

$|A^{-1}| = -\dfrac{1}{2} = \dfrac{1}{|A|}$. □

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ, 2, 2, [1, 2, 3, 4])
Ai=A.inverse()
print "det(A)=", A.det()
print "det(A^(-1))=", Ai.det()
```

```
det(A)= -2
det(A^(-1))= -1/2
```
■

[19th International Linear Algebra Conference(Sungkyunkwan University, 2014)]
http://www.ilas2014.org/

Photos and Movie: http://matrix.skku.ac.kr/2014-Album/ILAS-2014/

4.2 Cofactor Expansion and Applications of the Determinants

- Reference Video: http://youtu.be/XPCD0ZYoH5I, http://youtu.be/m6l2my6pSwY
- Practice Site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-5-Sec-4-2.html

In this section, we introduce a method which is convenient to compute the determinant as well as important in theory. Moreover, by applying this method, we introduce an easier formula to compute the inverse of a matrix.

Definition [Minor and cofactor]

We denote a submatrix, by removing the ith row and jth column of a given square matrix $A=[a_{ij}]$, as $A(i|j)$. We call its determinant $M_{ij}=\det A(i|j)$ as **minor** of A for a_{ij}. We also call $A_{ij}=(-1)^{i+j}|A(i|j)|=(-1)^{i+j}M_{ij}$ as **cofactor** of A for a_{ij}.

Example 1

For given matrix $A=\begin{bmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{bmatrix}$, find the minor and cofactor of A for a_{11}.

Solution The minor of A for a_{11} is $M_{11}=\det A(1|1)=\begin{vmatrix} 5 & 6 \\ 4 & 8 \end{vmatrix}=16$ and the cofactor of A for a_{11} is $A_{11}=(-1)^{1+1}M_{11}=16$. ■

Definition [Adjoint matrix]

Let A_{ij} be a cofactor of $n\times n$ matrix $A=[a_{ij}]$ for a_{ij}. The matrix $[A_{ij}]^T$ is called an **adjoint matrix** of A and is denoted by adj$A$. That is,

$$\text{adj } A=\begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}=\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}^T=[A_{ij}]^T$$

Example 2 Find $\operatorname{adj} A$ of the following matrix.

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

Solution

Note the cofactor of A for each element is as follows.

$$A_{11} = (-1)^{1+1}\begin{vmatrix} 6 & 2 \\ 0 & -3 \end{vmatrix} = -18 \quad A_{12} = (-1)^{1+2}\begin{vmatrix} 5 & 2 \\ 1 & -3 \end{vmatrix} = 17$$

$$A_{13} = (-1)^{1+3}\begin{vmatrix} 5 & 6 \\ 1 & -0 \end{vmatrix} = -6 \quad A_{21} = (-1)^{2+1}\begin{vmatrix} -2 & 1 \\ 0 & -3 \end{vmatrix} = -6$$

$$A_{22} = (-1)^{2+2}\begin{vmatrix} 3 & 1 \\ 1 & -3 \end{vmatrix} = -10 \quad A_{23} = (-1)^{2+3}\begin{vmatrix} 3 & -2 \\ 1 & 0 \end{vmatrix} = -2$$

$$A_{31} = (-1)^{3+1}\begin{vmatrix} -2 & 1 \\ 6 & 2 \end{vmatrix} = -10 \quad A_{32} = (-1)^{3+2}\begin{vmatrix} 3 & 1 \\ 5 & 2 \end{vmatrix} = -1$$

$$A_{33} = (-1)^{3+3}\begin{vmatrix} 3 & -2 \\ 5 & 6 \end{vmatrix} = 28$$

Therefore, $\operatorname{adj} A = \begin{bmatrix} -18 & -6 & -10 \\ 17 & -10 & -1 \\ -6 & -2 & 28 \end{bmatrix}$ □

- http://matrix.skku.ac.kr/RPG_English/4-MA-adjoint.html

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ, 3, 3, [3, -2, 1, 5, 6, 2, 1, 0, -3])
print A.adjoint()                              # adjoint matrix
```

```
[-18  -6 -10]
[ 17 -10  -1]
[ -6  -2  28]
```
■

Cofactor expansion

The determinant of 3×3 matrix $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ can be expanded as follows.

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

(Expand around the first column)

This is known as (Laplace) cofactor expansion of A around the first column.

Cofactor expansion works for any column and any row.

For any 3×3 matrix $A=[a_{ij}]$, the following identity holds. That is,

$$A\cdot\ \text{adj}\,A=\begin{bmatrix}\sum_{k=1}^{3}a_{1k}A_{1k} & \sum_{k=1}^{3}a_{1k}A_{2k} & \sum_{k=1}^{3}a_{1k}A_{3k}\\ \sum_{k=1}^{3}a_{2k}A_{1k} & \sum_{k=1}^{3}a_{2k}A_{2k} & \sum_{k=1}^{3}a_{2k}A_{3k}\\ \sum_{k=1}^{3}a_{3k}A_{1k} & \sum_{k=1}^{3}a_{3k}A_{2k} & \sum_{k=1}^{3}a_{3k}A_{3k}\end{bmatrix}=\begin{bmatrix}|A| & 0 & 0\\ 0 & |A| & 0\\ 0 & 0 & |A|\end{bmatrix}=|A|\cdot\ I_3\ .$$

Which shows $\sum_{k=1}^{3}a_{ik}A_{jk}=\begin{cases}|A| & (i=j)\\ 0 & (i\neq j)\end{cases}$.

Read: http://nptel.ac.in/courses/122104018/node29.html

For the previous example Example 2,

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ, 3, 3, [3, -2, 1, 5, 6, 2, 1, 0, -3])
print "det(A)=", A.det()
print "A*adj(A)="
print A*A.adjoint()
```

```
det(A)= -94
A*adj(A)=
[-94   0   0]
[  0 -94   0]
[  0   0 -94]
```

■

Therefore the following holds.

Theorem 4.2.1 [Cofactor expansion]

Let A be a $n \times n$ matrix. For any i, j $(1 \leq i, j \leq n)$ the following holds.

$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}$ (cofactor expansion around ith row)

$|A| = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}$ (cofactor expansion around jth column)

- When computing the determinant, it is advantageous to use the cofactor expansion around the row (column) which has many zeros.

Example 3 Find the determinant of a given matrix by using the cofactor expansion.

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

Solution

Multiply (-2) to the 2nd row and add it to the 3rd row. Multiply (-3) to the 2nd row and add it to the both 1st and 4th row. Then

$$|A| = \begin{vmatrix} 0 & -1 & 1 & 3 \\ 1 & 2 & -1 & 1 \\ 0 & 0 & 3 & 3 \\ 0 & 1 & 8 & 0 \end{vmatrix}.$$

Now we cofactor expand around the first column,

$$|A| = 0 + (1)(-1)^{2+1} \begin{vmatrix} -1 & 1 & 3 \\ 0 & 3 & 3 \\ 1 & 8 & 0 \end{vmatrix} + 0 + 0$$

$$= (-1)\,[0 + 0 + 3 - 9 + 24 - 0] = -18.$$ ■

Theorem 4.2.2 [Inverse matrix by using the adjoint matrix]

Let A be an $n \times n$ invertible matrix, then the inverse matrix of A is

$A^{-1} = \dfrac{1}{|A|} \operatorname{adj} A$.

Example 4

From Example 2, find the inverse matrix of A.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(QQ, 3, 3, [3, -2, 1, 5, 6, 2, 1, 0, -3])
dA=A.det()                          # compute the determinant
adjA=A.adjoint()                    # compute adjoint matrix
print "(1/dA)*adjA="
print (1/dA)*adjA                   # compute inverse matrix
print
print "A^(-1)="
print A.inverse()                   # compare the results of inverse matrix
```

```
(1/dA)*adjA=
[  9/47    3/47    5/47]
[-17/94    5/47    1/94]
[  3/47    1/47 -14/47]

A^(-1)=
[  9/47    3/47    5/47]
[-17/94    5/47    1/94]
[  3/47    1/47 -14/47]
[  3/47    1/47 -14/47]
```

■

Augment the
Teaching of
Linear
Algebra through the use of
Software
Tools

[ATLAST project] http://www1.umassd.edu/SpecialPrograms/Atlast/

4.3 Cramer's Rule

- Reference video: http://youtu.be/OlmrmmWXuvU, http://youtu.be/m2NkOX7gE50
- Practice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-6-Sec-4-3.html

In this section, we introduce **Cramer's rule** which is very useful tool for solving a system of linear equations.

Cramer's rule can be applied to systems of linear equations with the same number of unknowns and the equations.

Theorem 4.3.1 [Cramer's Rule]

For a system of linear equations,

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

let A be a coefficient matrix, and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$. Then the system of linear equations can be written as $A\mathbf{x} = \mathbf{b}$. If $|A| \neq 0$, the system of linear equations has a unique solution as follows:

$$x_1 = \frac{|A_1|}{|A|},\ x_2 = \frac{|A_2|}{|A|},\ \ldots,\ x_n = \frac{|A_n|}{|A|}.$$

Where $A_j\ (j = 1, 2, \cdots, n)$ denotes the matrix A with jth column replaced by the vector $\mathbf{b}$.

Proof Since $|A| \neq 0$, A is invertible. Hence the system of linear equations $A\mathbf{x} = \mathbf{b}$ has a unique solution $\mathbf{x} = A^{-1}\mathbf{b}$. Since $A^{-1} = \frac{1}{|A|}\operatorname{adj} A$, we have

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \left(\frac{1}{|A|}\operatorname{adj} A\right)\mathbf{b} = \frac{1}{|A|}\begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1j} & A_{2j} & \cdots & A_{nj} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Observe the jth component of $\mathbf{x}$ is $x_j = \dfrac{b_1 A_{1j} + b_2 A_{2j} + \cdots + b_n A_{nj}}{|A|}$. Since

$$a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} = |A|,$$

if we denote A_j as a matrix A with jth column replaced by the vector $\mathbf{b}$, then we have

$$x_j = \frac{|A_j|}{|A|} \quad (j = 1, 2, \cdots, n).$$ ■

Example 1 Solve the following system of linear equations by Cramer's rule.

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

Solution

Let A be the coefficient matrix. Then

$$|A| = \begin{vmatrix} -2 & 3 & -1 \\ 1 & 2 & -1 \\ -2 & -1 & 1 \end{vmatrix} = -2, \quad |A_1| = \begin{vmatrix} 1 & 3 & -1 \\ 4 & 2 & -1 \\ -3 & -1 & 1 \end{vmatrix} = -4,$$

$$|A_2| = \begin{vmatrix} -2 & 1 & -1 \\ 1 & 4 & -1 \\ -2 & -3 & 1 \end{vmatrix} = -6, \quad |A_3| = \begin{vmatrix} -2 & 3 & 1 \\ 1 & 2 & 4 \\ -2 & -1 & -3 \end{vmatrix} = -8, \text{ and hence}$$

$$x_1 = \frac{|A_1|}{|A|} = \frac{-4}{-2} = 2, \quad x_2 = \frac{|A_2|}{|A|} = \frac{-6}{-2} = 3, \quad x_3 = \frac{|A_3|}{|A|} = \frac{-8}{-2} = 4.$$ ■

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A = matrix(3,3,[-2,3,-1,1,2,-1,-2,-1,1]);
A1 = matrix(3,3,[1, 3, -1,4, 2,-1,-3,-1,1]);
A2 = matrix(3,3,[-2,1, -1, 1,4,-1,-2,-3,1]);
A3 = matrix(3,3,[-2,3,1, 1, 2, 4, -2,-1,-3]);
print A.det()
print A1.det()
print A2.det()
print A3.det()
print "x =",  A1.det()/A.det()
```

```
print "y =",  A2.det()/A.det()
print "z =",  A3.det()/A.det()
```

```
-2
-4
-6
-8
x = 2
y = 3
z = 4
```
■

Example 2 Solve the following system of linear equations by Cramer's rule.

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

Solution

From **Example 1** $|A| = -2$, and each matrix A_1, A_2, A_3 has zeros column, $|A_1| = |A_2| = |A_3| = 0$. Hence, the solution is z $x_1 = x_2 = x_3 = 0$ ■

Theorem 4.3.2 [Equivalence Theorem for Invertible Matrix]

For an $n \times n$ matrix A, the following are equivalent.

(1) $\text{RREF}(A) = I_n$

(2) A is a product of elementary matrices.

(3) A is invertible.

(4) $\mathbf{0}$ is the unique solution to $A\mathbf{x} = \mathbf{0}$.

(5) $A\mathbf{x} = \mathbf{b}$ has a unique solution for any $\mathbf{b} \in \mathbb{R}^n$.

(6) The columns of A are linearly independent.

(7) The rows of A are linearly independent.

(8) $|A| \neq 0$

Note that there are more equivalent statements for the above theorem. For more equivalent statements, refer Theorem 7.4.9 in section 7.4.

4.4 *Application of Determinant

- Reference video: http://youtu.be/OImrmmWXuvU, http://youtu.be/KtkOH5M3_Lc
- Practice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-6-Sec-4-4.html

The concept of determinant was first introduced by Japanese Takakazu Seki-Kowa in 1683. The term determinant originated from the meaning of *determining* the existence of roots. It was Cauchy who used the term in modern concept in 1815. In this section, we introduce some geometric and algebraic applications among many other applications of the determinant.

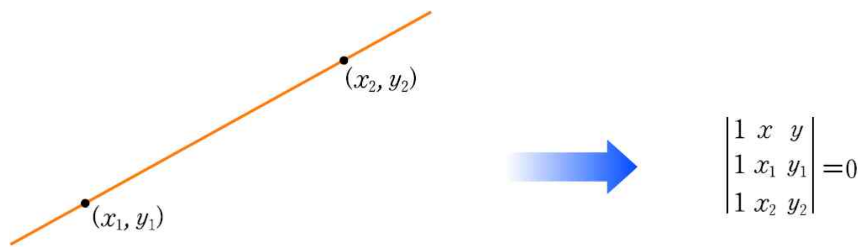
By using a determinant, we can easily find areas, volumes, equations of lines, equations of elliptic curves, or equations of plane. Also, the determinant of Vandermonde matrix connects between discrete data, which arise from statistical data and experimental labs, etc.

Example 1 Show that the equation of a line, which passes through two distinct points (x_1, y_1) and (x_2, y_2), is as follows.

$$\begin{vmatrix} 1 & x & y \\ 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \end{vmatrix} = 0$$

Solution

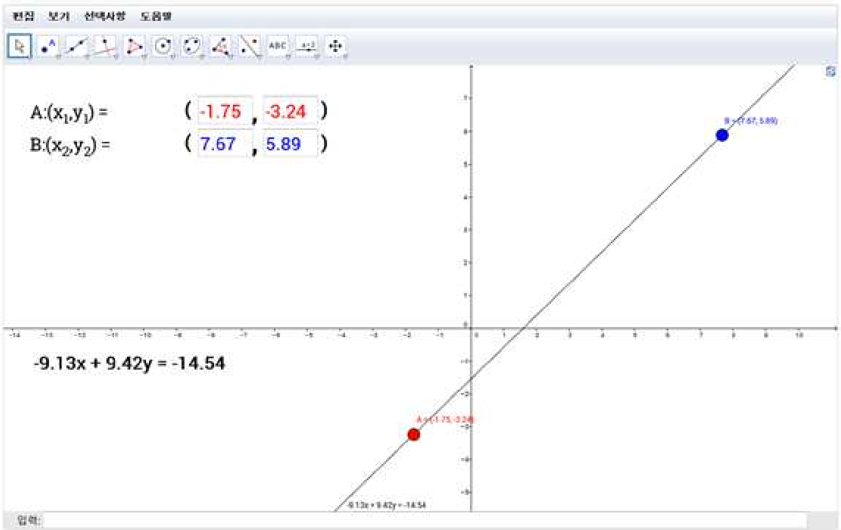
Note that the above equation is degree 1 for both x and y. As the equation holds by substituting $x = x_1$, $y = y_1$ and $x = x_2$, $y = y_2$ into the equation, the equation must be the equation of the line which passes through two given points.



■

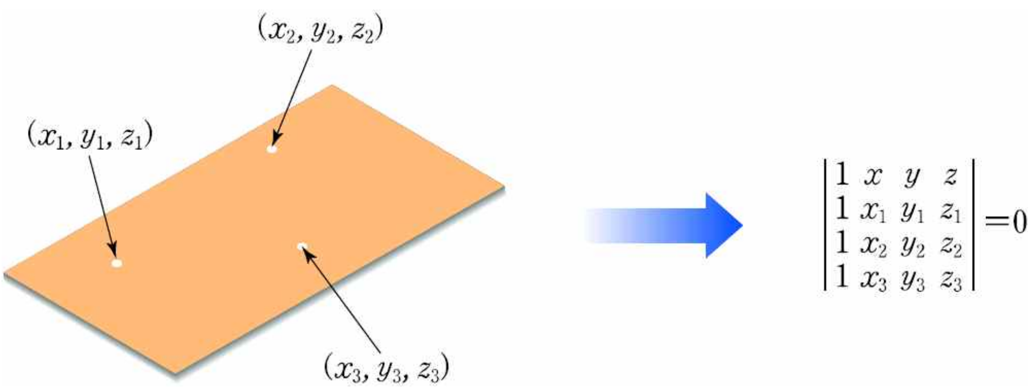
[Remark] Computer simulation

[An equation of a line which passes through two distinct points]
http://www.geogebratube.org/student/m9504



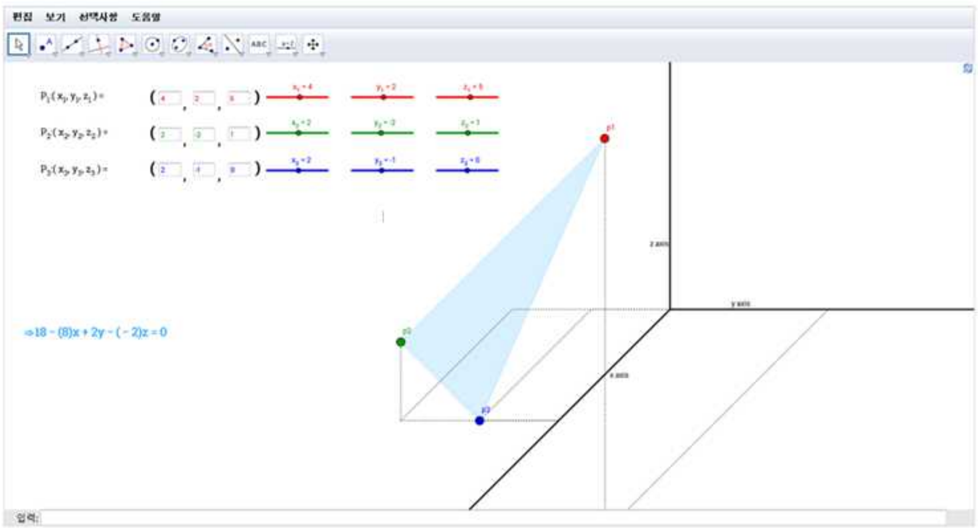
Similar to Example 1, an equation of a plane, which passes through three distinct points (x_1, y_1, z_1), (x_2, y_2, z_2), and (x_3, y_3, z_3), is as follows:

$$\begin{vmatrix} 1 & x & y & z \\ 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ 1 & x_3 & y_3 & z_3 \end{vmatrix} = 0 .$$



[Remark] **Computer simulation**

[An equation of a plane which passes through three distinct points]
http://www.geogebratube.org/student/m56430

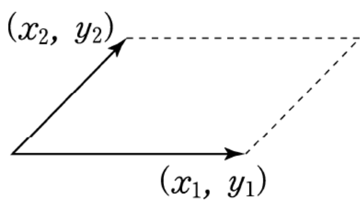


- Consider an arbitrary non-singular square matrix $A \in M_n$. Let $A^{(i)}$ be the ith column and

$$P(A) = \left\{ \sum_{i=1}^{n} t_i A^{(i)} \ : 0 \le \ t_i \le \ 1, \ i = 1, \ 2, \ \cdots, \ n \right\}.$$

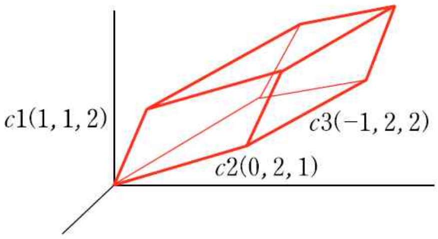
For the case $n = 2$ is a **parallelogram**, and for the case $n \ge 3$ is a generalized **parallelepiped**.

- Parallelogram can be expressed by adding two vectors as follows.



The area of the above parallelogram is $|x_1 y_2 - x_2 y_1|$ which is the same as the absolute value of $\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}$. Similarly, a parallelepiped is generated by three vectors which do not lie on the same plane. Let matrix A's columns consist by these three vectors. Then the volume of the parallelepiped is absolute value of $\det(A)$.

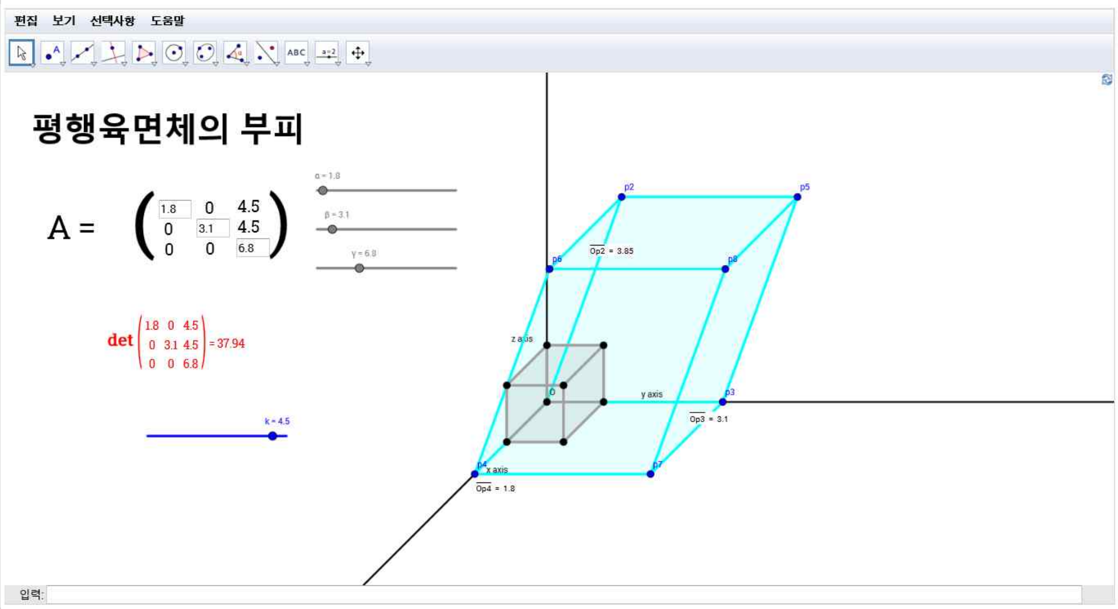
$$A = \begin{array}{c} \left[\begin{array}{ccc} \boxed{\begin{matrix}1\\1\\2\end{matrix}} & \boxed{\begin{matrix}0\\2\\1\end{matrix}} & \boxed{\begin{matrix}-1\\2\\2\end{matrix}} \end{array}\right] \\ \quad c1 \quad\; c2 \quad\; c3 \end{array}$$



$$\det A = \begin{vmatrix} 1 & 0 & -1 \\ 1 & 2 & 2 \\ 2 & 1 & 2 \end{vmatrix} = 5$$

[Remark] Computer simulation

[Volume of parallelepiped]
http://www.geogebratube.org/student/m57553



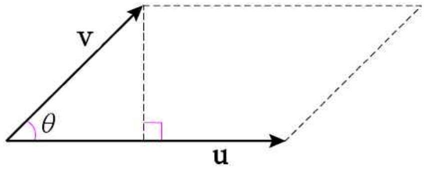
Theorem 4.4.1

(1) Let A be an 2×2 matrix. The area of parallelogram generated by two column vectors is $|\det(A)|$.

(2) Let A be an 3×3 matrix. The volume of parallelepiped generated by three column vectors is $|\det(A)|$.

(3) The area of parallelogram generated by two vectors $\mathbf{u}, \mathbf{v} \in R^n$, is $\sqrt{\det A^T A}$, where $A = [\mathbf{u} : \mathbf{v}]$.

Proof We will prove only (3).



Note that $|\mathbf{u}\cdot\ \mathbf{v}|^2 = ||\mathbf{u}||^2||\mathbf{v}||^2\cos^2\theta$.

Also, the area of parallelogram is $||\mathbf{u}||\,||\mathbf{v}||\sin\theta$. Now, the determinant

$$\begin{aligned}\det A^TA = \det\begin{bmatrix}\mathbf{u}^T\mathbf{u} & \mathbf{u}^T\mathbf{v}\\ \mathbf{v}^T\mathbf{u} & \mathbf{v}^T\mathbf{v}\end{bmatrix} &= ||\mathbf{u}||^2||\mathbf{v}||^2 - \mathbf{u}^T\mathbf{v}\mathbf{v}^T\mathbf{u}\\ &= ||\mathbf{u}||^2||\mathbf{v}||^2 - (\mathbf{v}^T\mathbf{u})^T(\mathbf{v}^T\mathbf{u}) = ||\mathbf{u}||^2||\mathbf{v}||^2 - |\mathbf{u}\cdot\ \mathbf{v}|^2\\ &= ||\mathbf{u}||^2||\mathbf{v}||^2 - ||\mathbf{u}||^2||\mathbf{v}||^2\cos^2\theta\\ &= ||\mathbf{u}||^2||\mathbf{v}||^2(1-\cos^2\theta) = ||\mathbf{u}||^2||\mathbf{v}||^2\sin^2\theta \ \text{(square of base times height)}\end{aligned}$$

makes the square of the area of the parallelogram generated by $\mathbf{u}, \mathbf{v} \in R^n$. ■

Example 2 Show that the area of a triangle generated by three points (x_1, y_1), (x_2, y_2), (x_3, y_3) is as follows.

$$\frac{1}{2}\left|\det\begin{bmatrix}x_1 & y_1 & 1\\ x_2 & y_2 & 1\\ x_3 & y_3 & 1\end{bmatrix}\right|$$

Solution

As the area is not changing by parallel translation, the area of triangle generated by three given vectors are the same as half of the area of parallelogram generated by (x_2-x_1, y_2-y_1) and (x_3-x_1, y_3-y_1). Hence, by Theorem 4.1.1, we have

$$\frac{1}{2}\left|\det\begin{bmatrix}x_2-x_1 & y_2-y_1\\ x_3-x_1 & y_3-y_1\end{bmatrix}\right| = \frac{1}{2}\left|\det\begin{bmatrix}x_1 & y_1 & 1\\ x_2-x_1 & y_2-y_1 & 0\\ x_3-x_1 & y_3-y_1 & 0\end{bmatrix}\right| = \frac{1}{2}\left|\det\begin{bmatrix}x_1 & y_1 & 1\\ x_2 & y_2 & 1\\ x_3 & y_3 & 1\end{bmatrix}\right|.\ ■$$

Vandermonde matrix and the determinants

If there are n distinct points in the xy-plane with mutually distinct x coordinates, then there exist a unique polynomial which passes through all given points with degree $n-1$. Let's find the polynomial.

- Let $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ be n distinct points in the xy-plane with mutually distinct x coordinates. We want to find a polynomial of degree $n-1$ which passes through all given points. Let

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1} \text{ be such a polynomial.}$$

Since these n points satisfy the given polynomial, we have

$$\begin{aligned} a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_{n-1} x_1^{n-1} &= y_1 \\ a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_{n-1} x_2^{n-1} &= y_2 \\ &\vdots \\ a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_{n-1} x_n^{n-1} &= y_n \ . \end{aligned}$$

Moreover, as $x_1, x_2, \cdots, x_n$ are mutually distinct, the coefficient matrix has

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ & & \vdots & & \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix} \neq 0 \ .$$

This coefficient matrix V_n is called **Vandermonde matrix** with degree n. Now we introduce how to compute the determinant of Vandermonde matrix. For the case $n = 3$,

$$\det V_3 = \begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} = \det(V_3^T) = \begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & x_2 - x_1 & x_3 - x_1 \\ 0 & x_2(x_2 - x_1) & x_3(x_3 - x_1) \end{vmatrix}$$

$$\begin{aligned} &= \begin{vmatrix} x_2 - x_1 & x_3 - x_1 \\ x_2(x_2 - x_1) & x_3(x_3 - x_1) \end{vmatrix} = (x_2 - x_1)(x_3 - x_1) \begin{vmatrix} 1 & 1 \\ x_2 & x_3 \end{vmatrix} \\ &= (x_2 - x_1)(x_3 - x_1)(x_3 - x_2) \end{aligned}$$

$$\therefore \quad |V_3| = \prod_{1 \le i < j \le 3} (x_j - x_i).$$

- Similarly, as we illustrated in the above case, the determinant of Vandermonde matrix V_n with degree n is product of $(x_j - x_i)$ (with $i < j$). That is,

$$\det V_n = \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ & & \vdots & & \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1 \le i < j \le n} (x_j - x_i) \ .$$

[Reference] http://www.proofwiki.org/wiki/Vandermonde_Determinant

Example 3 Find a polynomial that passes through the points (39, 34), (99, 47), (38, 58) by using a Vandermonde matrix.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

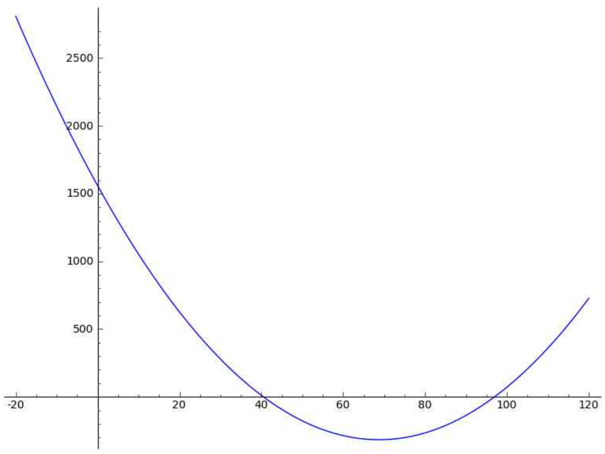
```
def Vandermonde_matrix(x_list):            # generate Vandermonde matrix
    n=len(x_list)
    A=matrix(RDF, n, n, [[z^i for i in range(n)] for z in x_list])
    return A

x_list=[39, 99, 38]                         # x coordinate
V=Vandermonde_matrix(x_list)
y_list=vector([34, 47, 58])                 # y coordinate
print "V="
print V
print
print "x=", V.solve_right(y_list)
```

```
V=
[    1.0    39.0  1521.0]
[    1.0    99.0  9801.0]
[    1.0    38.0  1444.0]

x= (1558.34590164, -54.568579235, 0.396994535519)
```

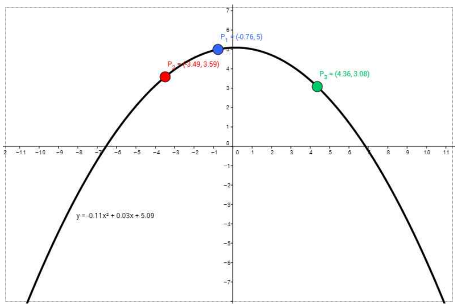
```
p=0.396994535519*x^2 -54.568579235*x + 1558.34590164
plot(p, (x, -20, 120))                # plot the graph
```

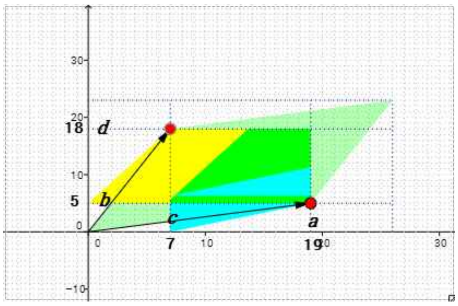


[Remark] Computer Simulation

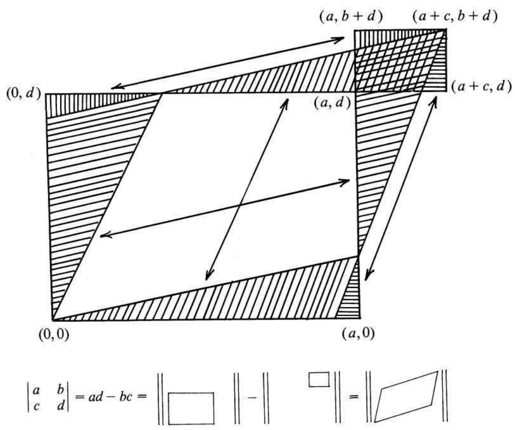
[Curve Fitting] http://www.geogebratube.org/student/m9911

[Area of parallelogram] http://www.geogebratube.org/student/m113





Proof without words:
A 2×2 determinant is the area of a parallelogram



[Solomon W. Golomb(Mathematics Magazine, March 1985)]

4.5 Eigenvalues and Eigenvectors

- Reference video: http://youtu.be/OlmrmmWXuvU, http://youtu.be/96Brbkx1cQ4
- Practice site: http://matrix.skku.ac.kr/knou-knowls/CLA-Week-6-Sec-4-5.html

For an $n \times n$ matrix A and a vector $\mathbf{x} \in R^n$, $A\mathbf{x}$ is a vector in R^n. One of the important questions in applied problems is "Is there any nonzero vector $\mathbf{x}$, which makes both $A\mathbf{x}$ and $\mathbf{x}$ parallel?" If such a vector exists, then it is called an eigenvector and it plays many important roles in linear transformation. In this section, we introduce eigenvectors and eigenvalues.

Definition [Eigenvalues and Eigenvectors]

Let A be an $n \times n$ matrix. For nonzero vector $\mathbf{x} \in R^n$, if there exist a scalar λ which satisfies $A\mathbf{x} = \lambda\mathbf{x}$, then λ is called an **eigenvalue** of A, and $\mathbf{x}$ is called an **eigenvector** of A corresponding to λ.

Example 1

Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Then $A\mathbf{x} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 3\mathbf{x}$. Hence 3 is an eigenvalue of A, and $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to 3.
■

Example 2

Since $I_n\mathbf{x} = 1\mathbf{x}$, for any $\mathbf{x} \in R^n$, the only eigenvalue of identity matrix I_n is $\lambda = 1$. Also, any nonzero vector $\mathbf{x} \in R^n$ is an eigenvector of I_n corresponding to 1. ■

- If $\mathbf{x} \in R^n$ is an eigenvector of A corresponding to λ, then $k\mathbf{x}$ is also an eigenvector of A corresponding to λ for any nonzero scalar k.

$$A\mathbf{x} = \lambda\mathbf{x} \;\Rightarrow\; A(k\mathbf{x}) = k(A\mathbf{x}) = k(\lambda\mathbf{x}) = \lambda(k\mathbf{x}).$$

General method to find eigenvalues

Since

$$A\mathrm{x} = \lambda\mathrm{x} \Leftrightarrow A\mathrm{x} = \lambda I_n\mathrm{x} \Leftrightarrow (\lambda I_n - A)\mathrm{x} = \mathbf{0}$$

and $\mathrm{x} \neq \mathbf{0}$, the system of linear equations $(\lambda I_n - A)\mathrm{x} = \mathbf{0}$ should have nonzero solution. Therefore, the characteristic equation, $|\lambda I_n - A| = 0$ should hold. $f_A(\lambda) = |\lambda I - A|$ is called the characteristic polynomial.

Theorem 4.5.1

Let A be $n \times n$ matrix and λ is a scalar, then the following statements are equivalent:

(1) λ is an eigenvalue of A.
(2) λ is a solution of the characteristic equation $\det(\lambda I_n - A) = 0$.
(3) System of linear equations $(\lambda I_n - A)\mathrm{x} = \mathbf{0}$ has a nontrivial solution.

Example 3 Find all eigenvalues and corresponding eigenvectors of $A = \begin{bmatrix} 5 & -6 \\ 2 & -2 \end{bmatrix}$.

Solution

If $\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ satisfies $A\mathrm{x} = \lambda\mathrm{x}$. Then,

$$\begin{bmatrix} 5 & -6 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Leftrightarrow \begin{matrix} 5x_1 - 6x_2 = \lambda x_1 \\ 2x_1 - 2x_2 = \lambda x_2 \end{matrix} \Leftrightarrow \begin{matrix} (\lambda - 5)x_1 + 6x_2 = 0 \\ -2x_1 + (\lambda + 2)x_2 = 0 \end{matrix} \qquad (1)$$

However, as mentioned above, this system of linear equations should have nontrivial (nonzero) solution. Hence,

$$\begin{vmatrix} \lambda - 5 & 6 \\ -2 & \lambda + 2 \end{vmatrix} = 0 \Leftrightarrow \lambda^2 - 3\lambda + 2 = 0 \Leftrightarrow (\lambda - 1)(\lambda - 2) = 0$$

$\therefore\ \lambda = 1,\ 2$

① Let's find an eigenvector corresponding to $\lambda_1 = 1$.

From (1), $\begin{matrix} -4x_1 + 6x_2 = 0 \\ -2x_1 + 3x_2 = 0 \end{matrix} \Leftrightarrow x_1 - \dfrac{3}{2}x_2 = 0$

$\therefore\ \mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3s \\ 2s \end{bmatrix} = s\begin{bmatrix} 3 \\ 2 \end{bmatrix}\ (s \in R \setminus \{0\})$

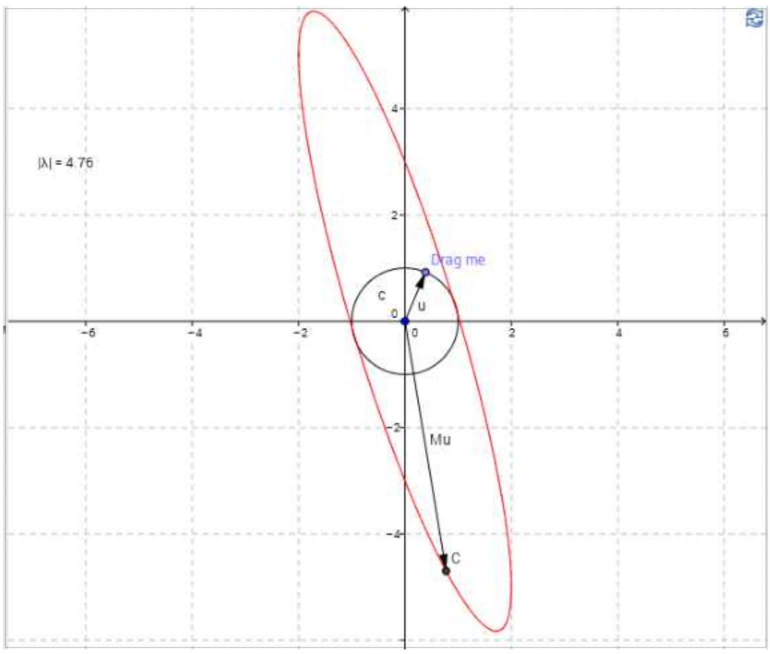
② Let's find an eigenvector corresponding to $\lambda_2 = 2$.

From (1), $\begin{array}{l} -3x_1 + 6x_2 = 0 \\ -2x_1 + 4x_2 = 0 \end{array} \Leftrightarrow \ x_1 - 2x_2 = 0$

$\therefore \ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2t \\ t \end{bmatrix} = t\begin{bmatrix} 2 \\ 1 \end{bmatrix} \ (t \in R \setminus \{0\})$ ■

[Remark] Computer simulation

[Visualize the eigenvalues and eigenvectors]
http://www.geogebratube.org/student/b73259#material/11114



- Do eigenvalues exist for any square matrix?

Theorem 4.5.2 [Fundamental Theorem of Algebra]

For any real (or complex) coefficient polynomial with degree n

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

has n roots $x_1, x_2, \ldots, x_n$, that is, $p(x_i) = 0$ for $i = 1, \ldots, n$, on the complex plane.

- That is, a real square matrix A with degree n always has n eigenvalues in complex domain. However, in this textbook we have limited the scalar as real numbers, and hence there is no eigenvalues means there is no *real* eigenvalues.

Example **4**

Find eigenvalues and eigenvectors of a matrix $A = \begin{bmatrix} 1 & -3 \\ -3 & 1 \end{bmatrix}$.

- http://matrix.skku.ac.kr/RPG_English/4-BN-char_ploy.html

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

① Characteristic equation of A

```
A=matrix(QQ, 2, 2, [1, -3, -3, 1])      # input A
print A.charpoly()                      # characteristic equation of A
```

```
x^2 - 2*x - 8
```

② Hence the eigenvalues are as follows.

```
solve(x^2 - 2*x - 8==0, x)
```

```
[x == -2, x == 4]
```

③ We can find the eigenvalues directly by using the built in command.

```
A.eigenvalues()                 # eigenvalues of A
```

```
[4, -2]
```

④ In order to find eigenvector for $\lambda = -2$, solve $(\lambda I_2 - A)\mathbf{x} = \mathbf{0}$.

```
(-2*identity_matrix(2)-A).echelon_form() # consider only coefficient matrix
```

```
[ 1 -1]
[ 0  0]
```

=> $x_1 - x_2 = 0$ ⇒ $\mathbf{x} = \begin{bmatrix} s \\ s \end{bmatrix} = s\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $(s \in R \setminus \{0\})$

⑤ In order to find eigenvector for $\lambda = 4$, solve $(\lambda I_2 - A)\mathbf{x} = \mathbf{0}$.

```
(4*identity_matrix(2)-A).echelon_form()   # consider only coefficient matrix
```

```
[ 1  1]
[ 0  0]
```

Hence, $x_1 + x_2 = 0$ ⇒ $\mathbf{x} = \begin{bmatrix} -t \\ t \end{bmatrix} = t\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ $(t \in R \setminus \{0\})$

⑥ We can find the eigenvectors directly by using the built in command.

```
A.eigenvectors_right()
```

```
[(4, [(1, -1)], 1), (-2, [(1, 1)], 1)]
# [eigenvalues, eigenvectors(it may appear in different form), multiplicity]
```

■

Example 5

Find eigenvalues and eigenvectors of a matrix $A = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ 3 & -3 & 0 \end{bmatrix}$.

- http://matrix.skku.ac.kr/RPG_English/4-VT-eigenvalues.html

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

① Characteristic equation of A

```
A=matrix(QQ, 3, 3, [1, 2, 2, 1, 2, -1, 3, -3, 0])  # input A
print A.charpoly()                               # characteristic equation of A
```

```
x^3 - 3*x^2 - 9*x + 27
```

② Hence the eigenvalues are as follows.

```
solve(x^3 - 3*x^2 - 9*x + 27==0, x)
```

```
[x == -3, x == 3]
```

③ We can find the eigenvalues directly by using the built in command.

```
A.eigenvalues()              # eigenvalues of A
```

```
[-3, 3, 3]                   # shows root with multiplicity 2
```

④ In order to find eigenvector for $\lambda = -3$, solve $(\lambda I_3 - A)\mathbf{x} = \mathbf{0}$.

```
(-3*identity_matrix(3)-A).echelon_form() # consider only coefficient matrix
```

```
[   1    0  2/3]
[   0    1 -1/3]
[   0    0    0]
```

Hence, $\begin{aligned} x_1 + \frac{2}{3}x_3 &= 0 \\ x_2 - \frac{1}{3}x_3 &= 0 \end{aligned}$ $\Rightarrow$ $\mathbf{x} = \begin{bmatrix} -2r \\ r \\ 3r \end{bmatrix} = r\begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}$ $(r \in R \setminus \{0\})$

⑤ In order to find eigenvector for $\lambda = 3$, solve $(\lambda I_3 - A)\mathbf{x} = \mathbf{0}$.

```
(3*identity_matrix(3)-A).echelon_form()  # consider only coefficient matrix
```

```
[ 1 -1 -1]
[ 0  0  0]
[ 0  0  0]
```

Hence, $x_1 - x_2 - x_3 = 0$ $\Rightarrow$ $\mathbf{x} = \begin{bmatrix} s+t \\ s \\ t \end{bmatrix} = s\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

(s and t are real numbers not simultaneously become zero)

⑥ We can find the eigenvectors directly by using the built in command.

```
A.eigenvectors_right()
```

```
[(-3, [(1, -1/2, -3/2)], 1), (3, [(1, 0, 1),(0, 1, -1)], 2)]
#[eigenvalues, eigenvectors(it may appear in different form), multiplicity]
```

For a triangular matrix $T=[t_{ij}]$ with degree n, the main diagonal components of $\lambda I-T$ are $\lambda - t_{ii}(i=1, 2, \ldots, n)$. Therefore, the characteristic polynomial of T is $\det(\lambda I-T)=(\lambda-t_{11})(\lambda-t_{22})\cdots(\lambda-t_{nn})$, and hence the eigenvalues of the triangular matrix T are its main diagonal components, $t_{11}, t_{22}, \ldots, t_{nn}$.

Example 6 Find the characteristic polynomial and all the eigenvalues of triangular matrix $T=\begin{bmatrix} 2 & 0 & 0 \\ -1 & -1 & 0 \\ 0 & 1 & 3 \end{bmatrix}$.

Solution

As $\det(\lambda I-T)=(\lambda-2)(\lambda+1)(\lambda-3)$, T's eigenvalues are $-1, 2, 3$. ■

Definition [Eigenspace]

Let λ be an eigenvalue of $n\times n$ matrix A. Then the solution space of the system of linear equations $(\lambda I_n - A)\mathbf{x}=\mathbf{0}$ is called **eigenspace** of A corresponding to λ.

- That is, an eigenspace of A corresponding to λ is the set of all eigenvectors of A corresponding to λ and the zero vector, which is a subspace of R^n.

Example 7 From the given matrix A in Example 5, find eigenspaces of A corresponding to each eigenvalue $\lambda_1=3$ and $\lambda_2=-3$.

Solution

From the result of Example 5,

① if $\lambda=-3$, by solving $(\lambda I_3 - A)\mathbf{x}=\mathbf{0}$, we have

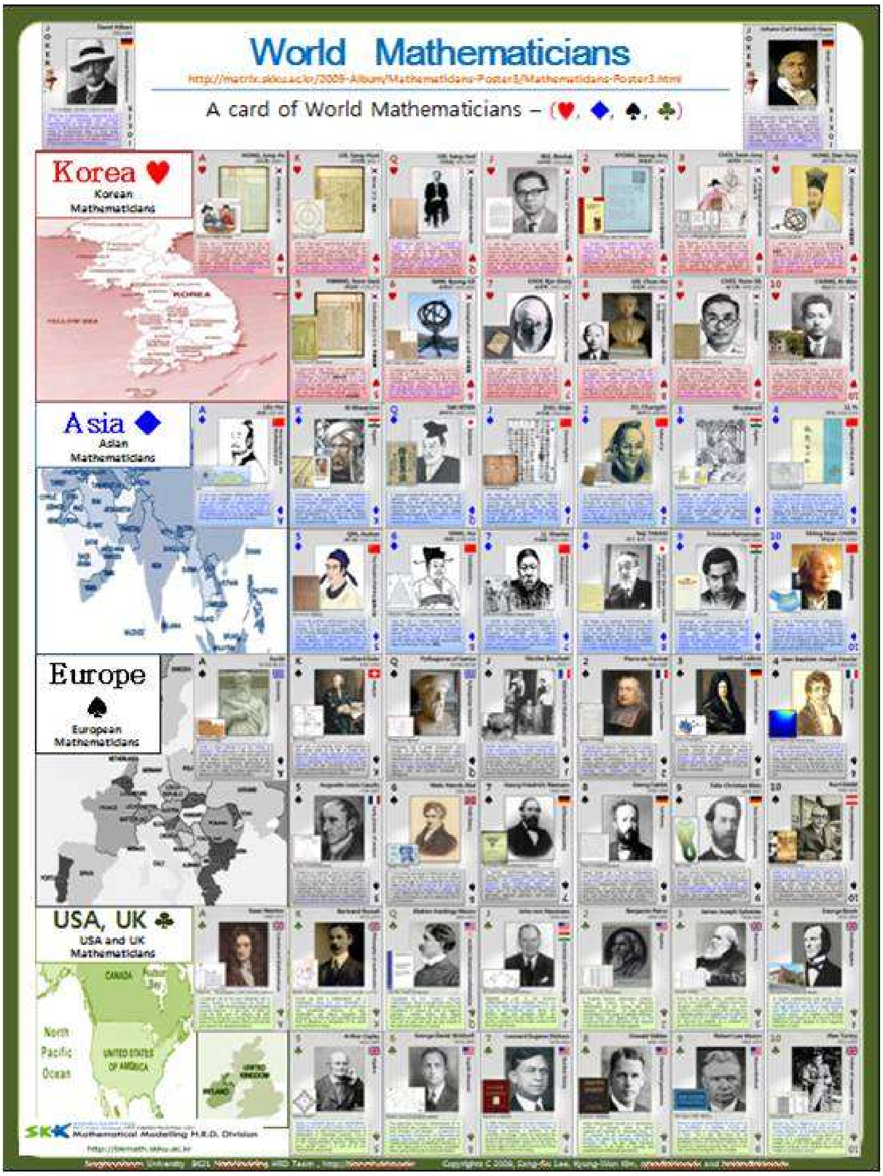
$$\begin{aligned} x_1+\frac{2}{3}x_3&=0 \\ x_2-\frac{1}{3}x_3&=0 \end{aligned} \quad \Rightarrow \quad \mathbf{x}=\begin{bmatrix} -2r \\ r \\ 3r \end{bmatrix}=r\begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} \quad (r\in R)$$

$$\therefore \quad W_1 = < \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} >$$

② When $\lambda = 3$, by solving $(\lambda I_3 - A)\mathbf{x} = \mathbf{0}$, we have

$$x_1 - x_2 - x_3 = 0 \quad \Rightarrow \quad \mathbf{x} = \begin{bmatrix} s+t \\ s \\ t \end{bmatrix} = s\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad (s,\ t \in R)$$

$$\therefore \quad W_2 = < \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} >$$ ■



http://matrix.skku.ac.kr/2009-Album/SKKU-Math-Card-F/SKKU-Math-Card-F.html

Chapter 4 Exercises

- http://matrix.skku.ac.kr/LA-Lab/index.htm
- http://matrix.skku.ac.kr/knou-knowls/cla-sage-reference.htm

Problem 1 Is permutation $(2\,0\,1\,3\,5\,10\,8\,7)$ of $S=\{0, 1, 2, 3, 5, 7, 8, 10\}$ even or odd?

Problem 2 Find the following determinants.

(1) $\det A = \begin{vmatrix} -3 & 5 & 1 \\ 1 & 0 & -2 \\ 0 & 3 & 0 \end{vmatrix}$ (2) $\det B = \begin{vmatrix} 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & -1 \\ 4 & 0 & 1 & 1 \\ 3 & 0 & 1 & 1 \end{vmatrix}$

Problem 3 Let A be $n \times n$ matrix and $|A| = -4$, find the followings.

(1) $|A^2|$

(2) $|A^{-1}|$

(3) $|2A|$

(4) $|(2A)^{-1}|$

Problem 4 For given matrices,

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

(1) show $|A| = |A^T|$.

(2) show $|AB| = |A||B|$.

(3) show $|A^{-1}| = \dfrac{1}{|A|}$.

Problem 5 For which x and y, the given matrix is invertible?

$$A = \begin{bmatrix} x^2 & 0 & 0 \\ y & x-1 & 0 \\ 1 & 2 & (y-1)(x-4) \end{bmatrix}$$

Problem 6 For given matrices,

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

(1) show $|A| = |A^T|$.

(2) show $|AB| = |A||B|$.

(3) show $|A^{-1}| = \dfrac{1}{|A|}$.

Problem 7 Find all cofactors of the following matrices.

(1) $A = \begin{bmatrix} 1 & 1 & 5 \\ 3 & -6 & 9 \\ 2 & 6 & 2 \end{bmatrix}$

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

Solution Sage :

```
A=matrix(QQ, 3, 3, [1, 1, 5, 0, 0, 0, 2, 6, 2])
print "adj A="
print A.adjoint()
B=matrix(QQ, 5, 5, [1, 2, 3, 4, 5, 0, 1, 0, 1, 0, -1, 1, -1, 1, -1, 0, 1, 2, 3, 4, -4, 2, -3, 1, 5])
print "adj B="
print B.adjoint()
```

adj A= [0 28 0] [0 -8 0] [0 -4 0]	adj B= [-18 36 -18 18 0] [14 -28 14 -14 0] [22 -44 22 -22 0] [-14 28 -14 14 0] [-4 8 -4 4 0]

■

Problem 8 Find the determinant of the matrix by cofactor expansion.

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

Problem 9 Find the adjoint matrix $\operatorname{adj} A$ of the matrix A from (Problem 8).

Problem 10 Find the inverse matrix of the given matrix by cofactor expansion.

$$(1)\ A=\begin{bmatrix} 1 & 0 & 1 \\ -1 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix} \qquad (2)\ A=\begin{bmatrix} 1 & 0 & 1 & 3 & 5 \\ -1 & 3 & 0 & 7 & 2 \\ 1 & 0 & 2 & 1 & 8 \\ 2 & -4 & 0 & 0 & 3 \\ -8 & 9 & 2 & 5 & 4 \end{bmatrix}$$

Solution (2) Sage : $A^{-1}=\dfrac{1}{|A|}\operatorname{adj} A$

```
A = matrix(QQ, 5, 5, [1, 0, 1, 3, 5, -1, 3, 0, 7, 2, 1, 0, 2, 1, 8, 2, -4, 0, 0, 3, -8, 9, 2, 5, 4])
dA = A.det()
adjA = A.adjoint()
print "(1/dA)*adjA="
print (1/dA)*adjA
```

```
(1/dA)*adjA=
[ -18/133    23/133      2/7  -48/133  -29/133]
[   -30/19     13/19        1    -4/19    -4/19]
[ 999/133 -412/133    -27/7 -129/133   80/133]
[ 164/133  -47/133     -5/7   -6/133   13/133]
[-268/133  106/133      8/7   39/133  -18/133]
```

$$\therefore A^{-1} = \begin{bmatrix} -\frac{18}{133} & \frac{23}{133} & \frac{2}{7} & -\frac{48}{133} & -\frac{29}{133} \\ -\frac{30}{19} & \frac{13}{19} & 1 & -\frac{4}{19} & -\frac{4}{19} \\ \frac{999}{133} & -\frac{412}{133} & -\frac{27}{7} & -\frac{129}{133} & \frac{80}{133} \\ \frac{164}{133} & -\frac{47}{133} & -\frac{5}{7} & -\frac{6}{133} & \frac{13}{133} \\ -\frac{268}{133} & \frac{106}{133} & \frac{8}{7} & \frac{39}{133} & -\frac{18}{133} \end{bmatrix}$$ ■

Problem 11 Solve the systems of linear equations by using the Cramer's rule.

(1) $\begin{cases} 3x - 3y - 2z = 3 \\ -x - 4y + 2z = 2 \\ 5x + 4y + \ z = 1 \end{cases}$

(2) $\begin{cases} x - y - z - w = 0 \\ -x - y + z + w = 2 \\ x + y - z + w = 1 \\ x + y + z + w = 1 \end{cases}$

(3) $\begin{cases} x_1 + 2x_2 + \ x_3 = 5 \\ 2x_1 + 2x_2 + \ x_3 = 6 \\ x_1 + 2x_2 + 3x_3 = 9 \end{cases}$

(4) $\begin{cases} x_1 + x_2 = 4 \\ x_2 + \ x_3 - 2x_4 = 1 \\ x_1 \ + 2x_3 + \ x_4 = 0 \\ x_1 + x_2 \ + \ x_4 = 0 \end{cases}$

Problem 12 Solve the following problems by using the determinant of Vandermonde matrix.

(1) Find the line equation which passes through the two points $(-1, 11)$ and $(2, -10)$.

(2) Find the coefficients a, b, c of parabolic equation $y = ax^2 + bx + c$ which passes through the three points $(1, 3)$, $(2, 3)$, and $(3, 5)$.

Problem 13 Solve the following problems by using the determinant.

(1) The area of a parallelogram which is generated by two sides connecting the origin and
each point $(4, 3)$ and $(7, 5)$.

(2) The volume of parallelepiped which is generated by three vectors, $(1, 0, 4)$, $(0, -2, 2)$, and $(3, 1, -1)$.

Problem 14 Find the eigenvalues and eigenvectors of the following matrices.

$$(1)\ A = \begin{bmatrix} 3 & 0 \\ -1 & -2 \end{bmatrix} \qquad (2)\ A = \begin{bmatrix} -3 & 0 & -2 & 8 \\ 0 & 1 & 4 & -2 \\ -4 & 10 & -1 & -2 \\ 6 & -4 & -2 & 3 \end{bmatrix}$$

Solution Sage : ① $\det(\lambda I - A) = 0$

```
A = matrix(QQ, 4, 4, [-3, 0, -2, 8, 0, 1, 4, -2, -4, 10, -1, -2, 6, -4, -2, 3])
print "character polynomial of A ="
print A.charpoly()
```

```
character polynomial of A =
x^4 - 118*x^2 - 168*x + 1485
```

② Eigenvalues

```
solve(x^4 - 118*x^2 - 168*x + 1485==0, x)
```

x == 11, x == -9, x == -5, x == 3]

④ Let x1, x2, x3 and x4 be the above eigenvalues.

```
A = matrix(QQ, 4, 4, [-3, 0, -2, 8, 0, 1, 4, -2, -4, 10, -1, -2, 6, -4, -2, 3])
x1 = 11
x2 = -9
x3 = -5
x4 = 3
print (x1*identity_matrix(4)-A).echelon_form()
print
print (x2*identity_matrix(4)-A).echelon_form()
print
print (x3*identity_matrix(4)-A).echelon_form()
print
print (x4*identity_matrix(4)-A).echelon_form()
```

```
[     1      0      0 -9/13]
[     0      1      0  7/13]
[     0      0      1 11/13]
[     0      0      0      0]

[ 1  0  0  2]
[ 0  1  0 -1]
[ 0  0  1  2]
[ 0  0  0  0]

[     1      0      0   7/5]
[     0      1      0   7/5]
[     0      0      1 -13/5]
[     0      0      0      0]

[ 1  0  0 -1]
[ 0  1  0 -1]
[ 0  0  1 -1]
[ 0  0  0  0]
```

⑤ Finding eigenvectors

```
A = matrix(QQ, 4, 4, [-3, 0, -2, 8, 0, 1, 4, -2, -4, 10, -1, -2, 6, -4, -2, 3])
print A.eigenvectors_right()
```

```
[(11, [(1, -7/9, -11/9, 13/9)], 1), (3, [(1, 1, 1, 1)], 1), (-5, [(1, 1, -13/7, -5/7)], 1), (-9, [(1, -1/2, 1, -1/2)],
1)]
```

∴ Eigenvectors corresponding to $\lambda_1=11$, $\lambda_2=-9$, $\lambda_3=-5$, $\lambda_4=3$ are $\mathbf{x}_1=\begin{bmatrix}9\\-7\\-11\\13\end{bmatrix}$, $\mathbf{x}_2=\begin{bmatrix}2\\-1\\2\\-1\end{bmatrix}$, $\mathbf{x}_3=\begin{bmatrix}7\\7\\-13\\-5\end{bmatrix}$, $\mathbf{x}_4=\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$. ■

Problem P1 Explain why $\det A=0$ for the following matrix A.

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

Problem P2 Show that for two square matrices A and B, if $A=P^{-1}BP$ for an invertible matrix P, then $|A|=|B|$.

Solution $|A|=|P^{-1}BP|=|P^{-1}||B||P|=|P^{-1}||P||B|=|P^{-1}P||B|=|I||B|=|B|$ ■

Problem P3 Simplify the following determinant.

$$\begin{vmatrix} (a+b)^2 & c^2 & c^2 \\ a^2 & (b+c)^2 & a^2 \\ b^2 & b^2 & (c+a)^2 \end{vmatrix}$$

Problem P4 For $n \times n$ matrix A, with $n > 1$, show the following identity.

$$\det(\operatorname{adj} A) = (\det A)^{n-1}$$

Problem P5 Let A be a 4×4 matrix, and assume that

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

(1) Find $\det(\operatorname{adj} A)$. Which relation does this value have with $\det(A)$?

(2) Find A.

Solution Sage :

```
adjA=matrix(QQ, 4, 4, [2, 0, 0, 0, 0, 2, 1, 0, 0, 4, 3, 2, 0, -2, -1, 2])
print adjA
print adjA.det()                          # |adj A|
B=(adjA.det())^(1/(4-1))                  # |A|^(n-1)=|adj A|
print B
C=(1/B)*adjA                              # A^(-1)
print C
print C.inverse()                         # A
D=matrix(QQ, 4, 4,[1,  0 , 0 , 0, 0 , 4 ,-1 , 1, 0, -6,  2, -2, 0 , 1 , 0,  1])
print D.adjoint()                         # adj A
```

Answers :

adjA=	det(adjA)=	A^(-1)=	A=	adjA=
[2 0 0 0] [0 2 1 0] [0 4 3 2] [0 -2 -1 2]	8 det(A)=2	[1 0 0 0] [0 1 1/2 0] [0 2 3/2 1] [0 -1 -1/2 1]	[1 0 0 0] [0 4 -1 1] [0 -6 2 -2] [0 1 0 1]	[2 0 0 0] [0 2 1 0] [0 4 3 2] [0 -2 -1 2]

■

Problem P6 By using the Cramer's rule, find the degree 3 polynomial $y=ax^3+bx^2+cx+d$ which passes through the following four points.

$$(0,\ 1),\ \ (1,\ -1),\ \ (2,\ -1),\ \ (3,\ 7)$$

Solution $\begin{cases} 1=d \\ -1=a+b+c+d \\ -1=8a+4b+2c+d \\ 7=27a+9b+3c+d \end{cases}$ => $\begin{cases} -2=a+b+c \\ -2=8a+4b+2c \\ 6=27a+9b+3c \end{cases}$

Sage :

```
A=matrix(3, 3, [1, 1, 1, 8, 4, 2, 27, 9, 3])
b=vector([-2, -2, 6])
Ai=A.inverse()
print "x=", Ai*b
print
print "x=", A.solve_right(b)
```

```
x= (1, -2, -1)
```

=> $a=1, b=-2, c=-1, d=1$ $\quad\therefore\ y=x^3-2x^2-x+1$ ■

Problem P7 Let the characteristic polynomial of matrix A be $p(\lambda)=(\lambda-1)(\lambda-2)^2$. Find eigenvalues of matrix A^2.

Problem P8 Find the eigenspaces of $A=\begin{bmatrix}1&2\\2&4\end{bmatrix}$ corresponding to each eigenvalue and show that they are orthogonal to each other in the plane.

Solution The eigenspace of A corresponding to $\lambda=0$ is E_1 = $<\begin{bmatrix}2\\-1\end{bmatrix}>$.

The eigenspace of A corresponding to $\lambda=5$ is $E_2 = <\begin{bmatrix}1\\2\end{bmatrix}>$.

Choose any $\mathbf{y}_1 = s\mathbf{x}_1$ and $\mathbf{y}_2 = t\mathbf{x}_2$ from E_1 and E_2, resp., then

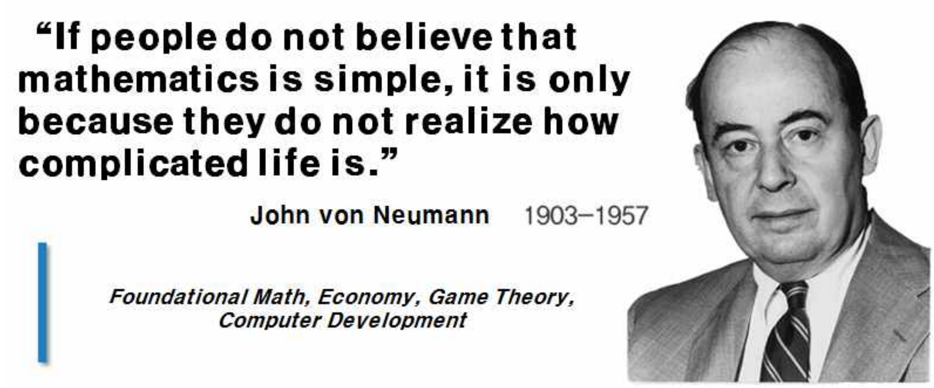
$<\mathbf{y}_1, \mathbf{y}_2> = < s\mathbf{x}_1, t\mathbf{x}_2> = st<\mathbf{x}_1, \mathbf{x}_2> = st(2\times1-1\times2)=0$.

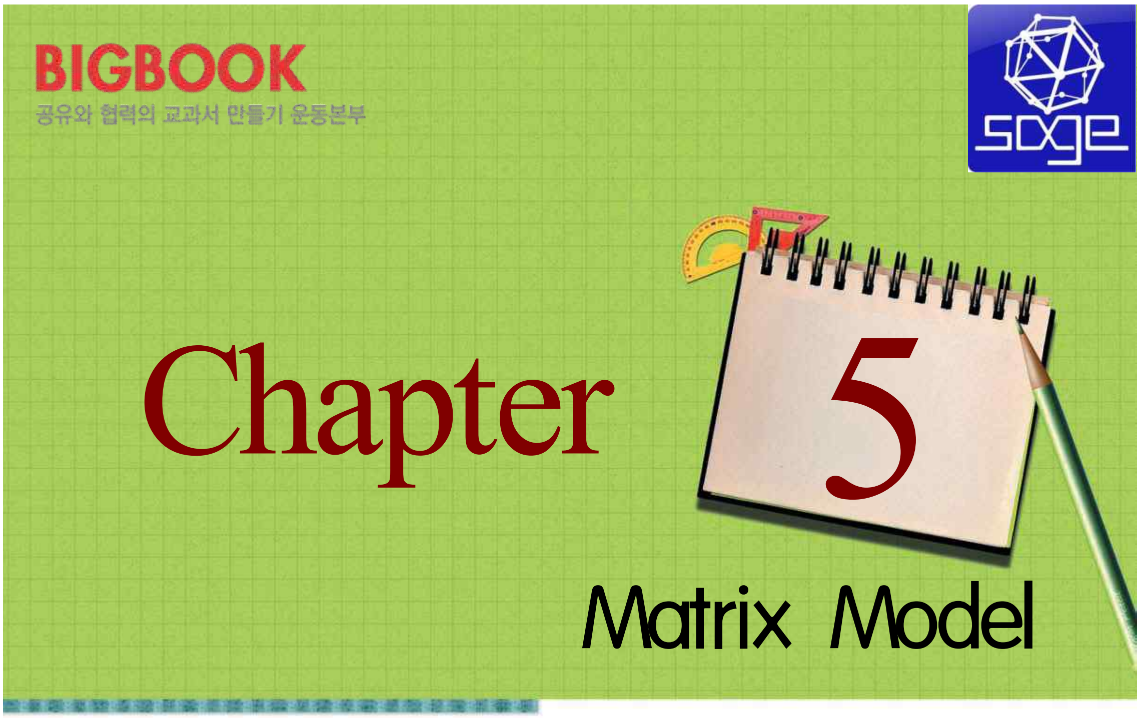
=> $\mathbf{y}_1$ and $\mathbf{y}_2$ are orthogonal.

$\therefore$ E_1 and E_2 are orthogonal to each other in the plane. ■

Problem P9 Find the characteristic polynomial of the following matrix. And find the roots of the polynomial by using the Sage.

$$A=\begin{bmatrix}1&1&2&1&1\\1&2&3&2&1\\2&3&1&2&1\\1&2&2&3&1\\1&1&1&1&7\end{bmatrix}$$





5.1 Lights out Game

5.2 Power Method

5.3 Linear Model (Google)

A mathematical model is a description of a system using mathematical concepts and language. The process of developing a mathematical model is called mathematical modeling. Mathematical models are used not only in the natural sciences (such as physics, biology, earth science, meteorology) and engineering disciplines (e.g. computer science, artificial intelligence), but also in the social sciences (such as economics, psychology, sociology and political science). Physicists, engineers, statisticians, operations research analysts and economists use mathematical models most extensively. A model may help to explain a system and to study the effects of different components and to make predictions about behaviour.

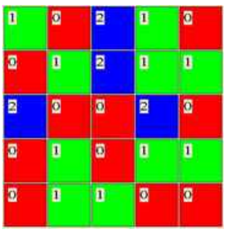
Mathematical models can take many forms, such as, dynamical systems, statistical models, differential equations, or game theoretic models. In this chapter, we **illustrate linear models**, which show how linear algebra can be used to solve the real world problems, and **review the content from previous chapters**.

http://matrix.skku.ac.kr/knou-knowls/cla-week-7.pdf

5.1 *Lights Out Game

- Reference video: http://youtu.be/_bS33lfa29s
- ractice site: http://matrix.skku.ac.kr/blackwhite2/blackwhite.html

http://matrix.skku.ac.kr/bljava/Test.html http://matrix.skku.ac.kr/Big-LA/Blackout.htm



The Blackout(Lights Out, Merlin's Magic square) Game, introduced in the official homepage of the popular movie `A Beautiful Mind', is a one-person strategy game that has recently gained popularity on handheld computing devices. An animated Macromedia Flash version of the game can be found from the official website of the 2001 movie `A Beautiful Mind'. In this section, we will introduce the question-and-answer process by one student that led to further development of this game, a purely linear algebraic solution and corresponding software.

Background of The Lights Out Puzzle

In my recent linear algebra class, we discussed the movie 'A Beautiful Mind', starring Russell Crowe as Nobel Laureate John F. Nash, Jr. (2001) specifically the scene where Nash was playing the game “Go” with one of his friends. Some of my students told me that they played 'the Blackout Puzzle' on the Korean official website of the movie.

http://www.abeautifulmind.com/

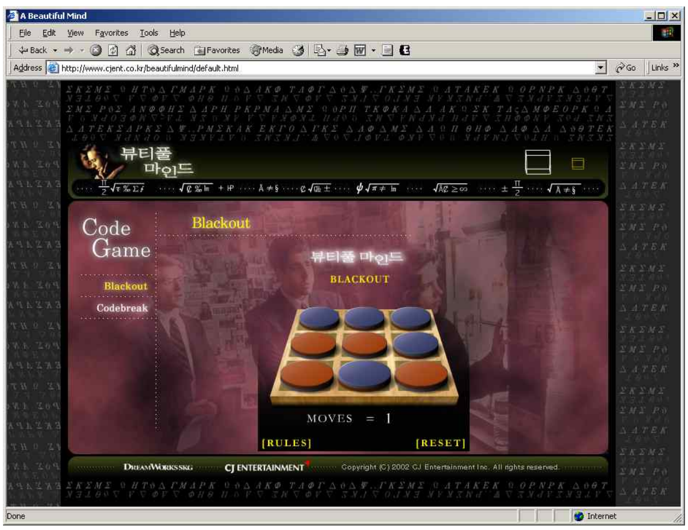


Figure 1: Blackout Game

One of my students asked me “Can we find an optimal solution for the game?” and, further, “Is there any possibility that we can not win the game if the given setting is fixed?” After a couple of days, one of my young students approached me with a potential solution. Together, we constructed a mathematical model of the Blackout Game and, utilizing this model, we were able to determine a solution to the original questions. What we found was that we can, in fact, always win the game, based on basic knowledge of linear algebra. At that time, the references about this game were limited, so we developed our own methods; it is these methods and results that will be explored in this section. Later, the following website was set up to further explain the puzzle and solutions:

(http://link.springer.com/article/10.1007/BF02896407 and http://matrix.skku.ac.kr/sglee/album/2004-ICME10SPF/ICME-10-July04.htm).

Introduction of blackout puzzle

A Blackout board is a grid of any size. Each square takes one of two colors black or white. (The diagram on the website as in Figure 1 used blue and red.) The player takes a turn by choosing any square, and the selected square and all squares that share an edge with it change their colors. The object of the game is to get all squares on the grid (tile) to be the same color - Black or White. When you click on a tile, the highlighted tile icons will change or “flip'' from their current state to the opposite state. Remember, the goal is to change all of the tile icons to black (or white).

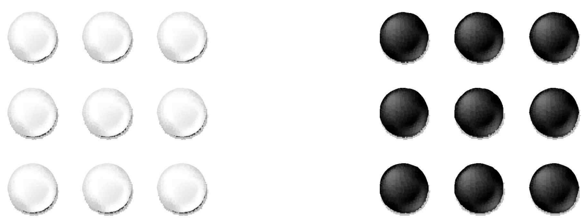


Figure 2: End of the Game (all squares having the same color)

How to solve any 3×3 game?

The following questions naturally come to mind:

[Q 1.] Is there any possibility that we can not win the game if the given setting is fixed?

[Q 2.] Given a winning pair (X_s, X_t), how many solutions are there? When is the solution unique?

[Q 3.] Can we make a program to give us an optimal solution (shortest sequence of moves)?

Note that here are $2\times2\times\cdots\times2=2^9=512$ patterns of 3×3 blackout grid. Among these 512 patterns, there are $9\times2=18$ patterns such that we can win the game with only one more click as follows. (Twice of the following basic 9 patterns as we can change all initial colors.)

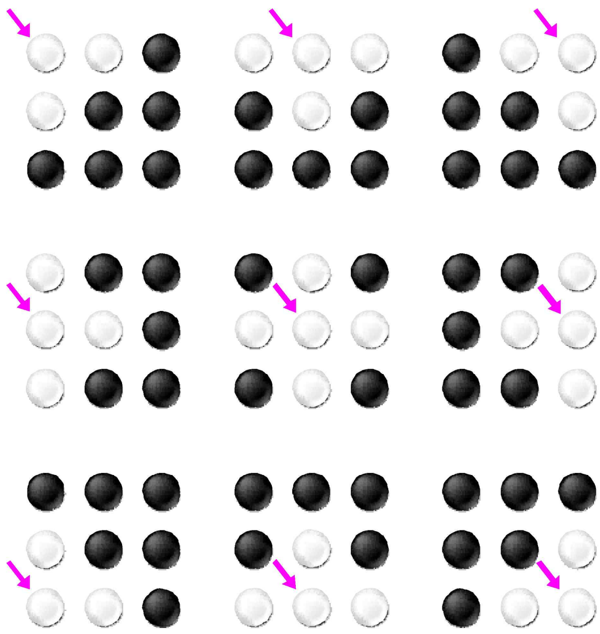


Figure 3

Adding some of the above to reach $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ or $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ (mod 2) is the goal of the game.

We checked several examples through trial-and-error to convince us of the answer to the first question regarding any given initial condition.

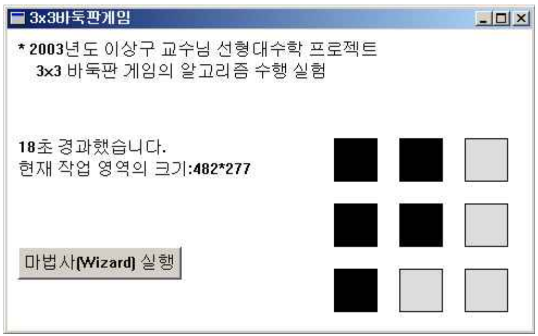
The figure 3 illustrates the shortest sequence of moves for resolving possible scenarios on a 3×3 board. Our approach to find a winning strategy was to recognize these 18 patterns in Figure 3.

Then, we tried to make a mathematical model of this game that the only actions we can perform are 9 clicks (since there are only 9 stones on the board). We assumed "the white stone ≡ 1 and black stone ≡ 0". Then, we classified effects of each action as an addition of one vector (or 3×3 matrix). Any series of our actions results in a linear combination of these vectors. We used modular 2 arithmetic to make the zero vector or all 1's vector (or matrix, resp.) to finish the game.

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

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

Thus, we now have the 9 vectors as shown above (in fact, twice the amount of them) to consider, which will end the game with just one more click. Assume the following initial condition, and the following 3 clicks make the entire board all white. Suppose we have 5 black(blue) stones and 4 white(red) stones in the board as below.



Then, the above condition can be denoted by the following matrix

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

Now, we choose some of 9 positions to act on it. This can be represented by

$$a\begin{bmatrix}1\,1\,0\\1\,0\,0\\0\,0\,0\end{bmatrix}+b\begin{bmatrix}1\,1\,1\\0\,1\,0\\0\,0\,0\end{bmatrix}+c\begin{bmatrix}0\,1\,1\\0\,0\,1\\0\,0\,0\end{bmatrix}+d\begin{bmatrix}1\,0\,0\\1\,1\,0\\1\,0\,0\end{bmatrix}+e\begin{bmatrix}0\,1\,0\\1\,1\,1\\0\,1\,0\end{bmatrix}+f\begin{bmatrix}0\,0\,1\\0\,1\,1\\0\,0\,1\end{bmatrix}+g\begin{bmatrix}0\,0\,0\\1\,0\,0\\1\,1\,0\end{bmatrix}+h\begin{bmatrix}0\,0\,0\\0\,1\,0\\1\,1\,1\end{bmatrix}+i\begin{bmatrix}0\,0\,0\\0\,0\,1\\0\,1\,1\end{bmatrix}$$

thus our problem is to find some a,b,c,d,e,f,g,h and i such that

$$a\underset{(1,1)}{\begin{bmatrix}1&1&0\\1&0&0\\0&0&0\end{bmatrix}}+b\underset{(1,2)}{\begin{bmatrix}1&1&1\\0&1&0\\0&0&0\end{bmatrix}}+c\underset{(1,3)}{\begin{bmatrix}0&1&1\\0&0&1\\0&0&0\end{bmatrix}}+d\underset{(2,1)}{\begin{bmatrix}1&0&0\\1&1&0\\1&0&0\end{bmatrix}}+e\underset{(2,2)}{\begin{bmatrix}0&1&0\\1&1&1\\0&1&0\end{bmatrix}}$$

$$+f\underset{(2,3)}{\begin{bmatrix}0&0&1\\0&1&1\\0&0&1\end{bmatrix}}+g\underset{(3,1)}{\begin{bmatrix}0&0&0\\1&0&0\\1&1&0\end{bmatrix}}+h\underset{(3,2)}{\begin{bmatrix}0&0&0\\0&1&0\\1&1&1\end{bmatrix}}+i\underset{(3,3)}{\begin{bmatrix}0&0&0\\0&0&1\\0&1&1\end{bmatrix}}+\underset{\text{Initial}}{\begin{bmatrix}0&0&1\\0&0&1\\0&1&1\end{bmatrix}}=\underset{\text{Final Goal}}{\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}}$$

$$\Rightarrow\quad a\underset{(1,1)}{\begin{bmatrix}1&1&0\\1&0&0\\0&0&0\end{bmatrix}}+b\underset{(1,2)}{\begin{bmatrix}1&1&1\\0&1&0\\0&0&0\end{bmatrix}}+c\underset{(1,3)}{\begin{bmatrix}0&1&1\\0&0&1\\0&0&0\end{bmatrix}}+d\underset{(2,1)}{\begin{bmatrix}1&0&0\\1&1&0\\1&0&0\end{bmatrix}}+e\underset{(2,2)}{\begin{bmatrix}0&1&0\\1&1&1\\0&1&0\end{bmatrix}}$$

$$+f\begin{bmatrix}0&0&1\\0&1&1\\0&0&1\end{bmatrix}+g\begin{bmatrix}0&0&0\\1&0&0\\1&1&0\end{bmatrix}+h\begin{bmatrix}0&0&0\\0&1&0\\1&1&1\end{bmatrix}+i\begin{bmatrix}0&0&0\\0&0&1\\0&1&1\end{bmatrix}=\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}-\begin{bmatrix}0&0&1\\0&0&1\\0&1&1\end{bmatrix}=-\mathrm{b}$$

(2,3) (3,1) (3,2) (3,3) Final Goal Initial

We now consider 3×3 matrix $\begin{bmatrix}1&1&0\\1&0&0\\0&0&0\end{bmatrix}$ as a 9×1 vector $\begin{bmatrix}1\\1\\0\\1\\0\\0\\0\\0\\0\end{bmatrix}$, then the above linar system of equation can be written as

$$A\mathrm{x}=-\mathrm{b}\ \Rightarrow\ A\begin{bmatrix}a\\b\\c\\d\\e\\f\\g\\h\\i\end{bmatrix}=-\begin{bmatrix}0\\0\\1\\0\\0\\1\\0\\1\\1\end{bmatrix}\ \text{where } A=\begin{bmatrix}1&1&0&1&0&0&0&0&0\\1&1&1&0&1&0&0&0&0\\0&1&1&0&0&1&0&0&0\\1&0&0&1&1&0&1&0&0\\0&1&0&1&1&1&0&1&0\\0&0&1&0&1&1&0&0&1\\0&0&0&1&0&0&1&1&0\\0&0&0&0&1&0&1&1&1\\0&0&0&0&0&1&0&1&1\end{bmatrix}$$

We can use any computational tool such as Sage and obtain

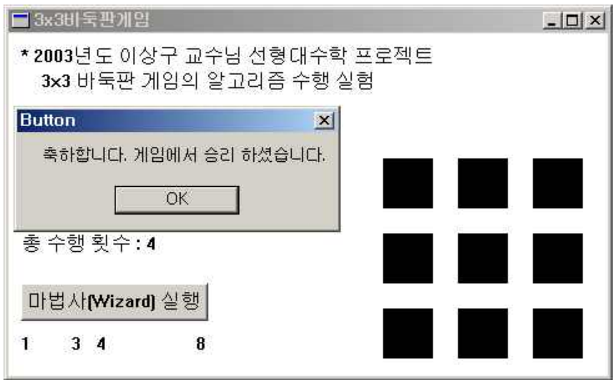
$$A^{-1}=\frac{1}{7}\begin{bmatrix}-1&4&-1&4&-2&-3&-1&-3&6\\4&-2&4&-2&1&-2&-3&5&-3\\-1&4&-1&-3&-2&4&6&-3&-1\\4&-2&-3&-2&1&5&4&-2&-3\\-2&1&-2&1&3&1&-2&1&-2\\-3&-2&4&5&1&-2&-3&-2&4\\-1&-3&6&4&-2&-3&-1&4&-1\\-3&5&-3&-2&1&-2&4&-2&4\\6&-3&-1&-3&-2&4&-1&4&-1\end{bmatrix}.$$

Then we have a system of linear equations to find $\mathrm{x}=[a\,b\,c\,d\,e\,f\,g\,h\,i]^T$. $A\mathrm{x}=-\mathrm{b}$ is a given (condition) matrix and j is a vector of all 1's. Then $\mathrm{RREF}(A)=I_9$ and $\mathrm{rank}A=9$. So the columns (rows) are linearly independent, and the system has a unique solution $\mathrm{x}=\left[\frac{1}{7}\ -\frac{4}{7}\ \frac{1}{7}\ \frac{3}{7}\ \frac{2}{7}\ -\frac{4}{7}\ -\frac{6}{7}\ \frac{3}{7}\ -\frac{6}{7}\right]^T$. Furthermore. this entire process can be done in Modular 2 arithmetic and $7\mathrm{x}$ = $[1\ -4\ 1\ 3\ 2\ -4\ -6\ 3\ -6]^T$. We only need 0 and 1 because clicking $2n+1$ times of one stone is the same as clicking once, and $2n$ clickings of one stone is the same as doing nothing. So, our answer for $A\mathrm{x}=-\mathrm{b}$, which is a real optimal winning strategy vector (matrix) $7\mathrm{x}\equiv[1\,0\,1\,1\,0\,0\,0\,1\,0]^T\ (\mathrm{mod}\ 2)$, is

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

This shows that if we click on positions $(1,1),(1,3),(2,1),(3,2)$, we will get all white stones on the board with only 4 clicks. With this idea, one of my students made a computer program in C++ based on this algorithm to determine an optimal winning strategy. Let $\mathbf{x}' \equiv \mathbf{x} \bmod 2$. Then $\mathbf{x}'$ is a real optimal winning strategy vector (matrix) which can be deduced from $\mathbf{x}$. Now, entries of $\mathbf{x}'$ are all 0 or 1 as is in real game situation and we can always find a (0,1) matrix as a real optimal winning strategy vector(matrix). We can download this program and run it from http://matrix.skku.ac.kr/sglee/blackout_win.exe. This software also verified our conjecture and showed the proof was valid.

In the following Figure, the command “(Wizard)” tells us “1 3 4 8,” which indicates which 4 stones we have to click to win. The number ''4'' shows we won with 4 clicks (MOVE).

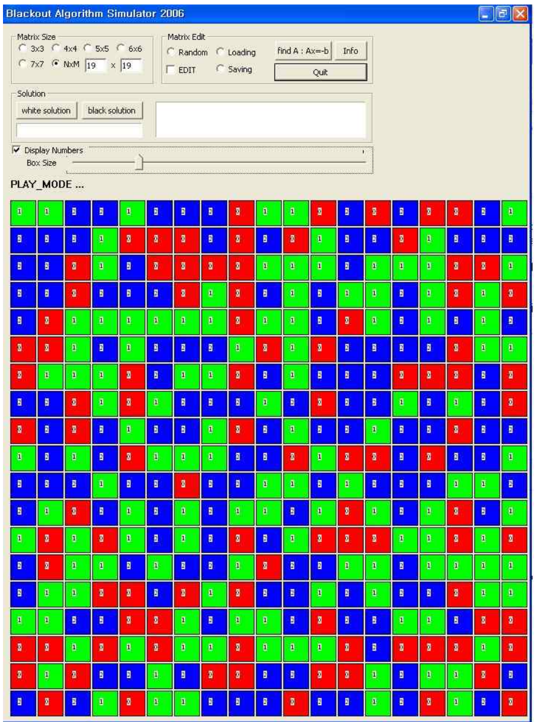


Teachers often think of "teaching" as a one-sided process, but this experience shows that teachers and creative students can work together to solve problems in a mathematically-stimulating, mutually beneficial way. This process can be adapted to resolve other real world problems using basic mathematical knowledge.

After answering our posed questions for the Blackout Game, we looked toward finding a relationship between the Blackout Game and automata theory. We started to introduce the concept of sigma-game and find the optimal strategy to win the Blackout game, as well as a condition to determine the irreversibility of this game in larger size boards- up to 19×19. We also verify our algorithm within a program made in C++.

The sigma-game is played on a directed graph G. We suppose that the vertices of G can have one of two different states, which are designated as 0 or 1. A configuration is an assignment of states to all the vertices, and a move in the game consists of the player's picking a vertex. The Blackout Game emulates the sigma-game on the nine point directed graph.

We could classify the reversibility as a direct calculation of the $N \times N$ block tridiagonal matrix of the blackout game of size n. In fact, for $n \leq 19$, there are irreversible cases when $n = 4,5,9,11,14,17,19$. Using Mathematica and our eigenvalue method, we can easily show the irreversibility. We could find a way to reach the goal even for some irreversible case if we give a restriction on the initial condition b. This complete our generalization of the Blackout game from 3×3 board to the full size Go board. Finally, the following Figure from our software shows our answer is accurate for larger size boards.



[A software of Blackout game on different sizes with 3 colors]

We made a mathematical model from the well-known Blackout game. Surprisingly, it turned out to be a pure linear algebra problem of finding the optimal solution of the game, and we generalized it to the full size Go board. We gave a mathematical proof and algorithm to solve it which can be extended to the study of sigma-automata theory.

More details on the blackout game can be obtained from the following links.

http://matrix.skku.ac.kr/sglee/blackout_win.exe
http://matrix.skku.ac.kr/sglee/blackout_win.zip
http://matrix.skku.ac.kr/bljava/Test.html
http://matrix.skku.ac.kr/2012-mm/lectures-2012/A3-blackout-paper-ENG.pdf
http://matrix.skku.ac.kr/2009/2009-MathModeling/lectures/week12.pdf

[References]

Duk-Sun Kim, Sang-Gu Lee, Faqir M. Bhatti, Fibonacci sequences and the winning conditions for the blackout game, *International Journal of Contemporary Mathematical Sciences*, Vol. 5, 2010, no. 30, 1471 - 1485.

S.-G. Lee, I.-P. Kim, J.-B. Park, and J.-M. Yang, Linear algebra algorithm for the optimal solution in the Blackout game, *J. of KSME Ser. A*, 43 (2004), 87-96.

H.-S. Park, Go Game With Heuristic Function, *Kyongpook Nat. Univ. Elec. Tech Jour.* 15, No.2 (1994), 35-43.

D. H. Pelletier, Merlin's magic square, *Amer. Math. Monthly*, 94 (1987) 143-150

K. Sutner, The sigma-game and cellular automata, *Amer. Math. Monthly*, 97 (1990), 24-34.

J. Uhl, W. Davis, Is the mathematics we do the mathematics we teach?, *Contemporary issues in mathematics education*, 36 (1999), 67-74, Berkeley: MSRI Publications, Linear algebraic approach on real sigma-game

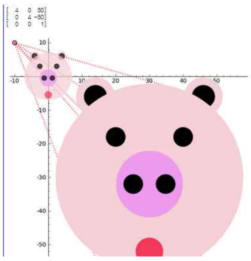
JAVA program by Universal Studios and DreamWorks, Movie: A Beautiful Mind, Blackout puzzle(2001), http://www.abeautifulmind.com/

3D Printing object 2

http://matrix.skku.ac.kr/2014-Album/2014-12-ICT-DIY/index.html

5.2 *Power Method

- Reference video: http://youtu.be/CLxjkZuNJXw
- Practice site: http://matrix.skku.ac.kr/2012-LAwithSage/interact/
  http://math1.skku.ac.kr/home/pub/1516/
  http://matrix.skku.ac.kr/SOCW-Math-Modelling.htm



In many matrix models of social behavior, the corresponding **maximum eigenvalue** gives adequate information to predict the model. Hence, often finding the maximum eigenvalue is enough to solve the corresponding problem. However, if the size of a matrix is significantly large, even with a computer, it is difficult to find all eigenvalues explicitly. Hence **for a large scale matrix, we look at a new method which finds *only* the maximum eigenvalue instead of finding all the eigenvalues.** This method, which harnesses the power of the matrix, is called **"Power Method"**. The first goal of this section is to explain how we can find **the maximum eigenvalue numerically**. The second goal is to show how this can be applied to the Google search engine.

We know that finding eigenvalues of an $n \times n$ real square matrix amounts to finding roots of its characteristic polynomial of degree n. However, for n large, finding the roots of n-th degree polynomial is not an easy task. Also finding numerical roots for a large degree polynomial is sensitive to rounding off errors.

In this article, we discuss numerical methods to approximate a largest or dominant eigenvalue of a matrix if exists. The dominant eigenvalues of a matrix have several applications in science, engineering and economics. Google uses it for page ranking the web pages and Twitter uses it to recommend users "WHO-TO-FOLLOW" (WTF).

Definition

Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of an $n \times n$ real matrix A. Then λ_1 is called a **dominant eigenvalue** of A if $|\lambda_1| > |\lambda_i|$ for all $i = 2, \ldots, n$.

The eigenvector corresponding to the dominant eigenvalue is called the **dominant eigenvector**.

[Remark] *Note that not every matrix possesses a dominant eigenvalue.*
For example, matrices

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

do not have dominant eigenvalues.

Power Method

Let A be an $n \times n$ real matrix. The power method is a numerical approach to find the dominant eigenvalue and the corresponding dominant eigenvector. We assume the following two conditions:

- The dominant eigenvalue is a real number and its absolute value is strictly greater than all the other eigenvalue.
- A is diagonalizable, in particular A has n linearly independent eigenvectors.

Let A have n linearly independent eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ and eigenvalues are orders as

$$|\lambda_1| < |\lambda_2| \leq |\lambda_3| \leq \cdots \leq |\lambda_n|.$$

Now we start with any nonzero vector $\mathbf{x}^{(0)} \in \mathbb{R}^n$and we continually multiply $\mathbf{x}^{(0)}$ by A which generates a sequence of vectors $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(k)}$, where

$$\mathbf{x}^{(k+1)} = A\mathbf{x}^{(k)},\ k = 0, 1, 2, \ldots$$

This implies $A^{(k)} = A\mathbf{x}^{(k-1)} = A^2\mathbf{x}^{(k-2)} = \cdots = A^k\mathbf{x}^{(0)}$.

Since A has n linearly independent eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$, there exist scalars $c_1, \ldots, c_n$such that

$$\mathbf{x}^{(0)} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n.$$

Multiplying both sides by A^k, we get

$$\begin{aligned}\mathbf{x}^{(k)} &= A^k\mathbf{x}^{(0)} = \sum_{i=1}^{n} c_i A^k \mathbf{x}_i \\ &= \sum_{i=1}^{n} c_i \lambda_i^k \mathbf{x}_i = \lambda_1^k\left[c_1\mathbf{x}_1 + \sum_{i=2}^{n} c_i\left(\frac{\lambda_i}{\lambda_1}\right)^k \mathbf{x}_i\right]\end{aligned}$$

Since $\lambda_1 > \lambda_i$ for all $i > 1$, the ration $\left|\frac{\lambda_i}{\lambda_1}\right| < 1$. Thus as $k \to \infty$, $\frac{\lambda_i}{\lambda_1} \to 0$. Hence $\frac{x^{(k)}}{\lambda_1^k} \to c_1\mathbf{x}_1$.

This leads to one of the very important method of finding the dominant eigenvalue and eigenvector, namely the "Power Method".

While applying the power method algorithm, we make sure that the largest component each of $\mathbf{x}^{(k)}$ is unity, in this case the component of $\mathbf{x}^{(k+1)} = A\mathbf{x}^{(k)}$ will have largest component of absolute value of λ.

Power Method Algorithm

[Step 1] Select the a vector $\mathbf{x}^{(0)}$ having largest component as 1.
[Step 2] Set $k = 0$.
[Step 3] Find $\mathbf{y}^{(k)} = A\mathbf{x}^{(k)}$.
[Step 4] Define c_k to be largest component in absolute value in the vector $\mathbf{x}^{(k)}$.
[Step 5] Define $\mathbf{x}^{(k+1)} = \frac{1}{c_k}\mathbf{y}^{(k)}$.
[Step 6] Check if the convergence criteria is met. Otherwise
[Step 7] Set $k = k+1$ and go the the step 3.

Example 1

Let us find the dominant eigenvalue and the corresponding eigenvector of the matrix $A = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}$, starting with $\mathbf{x}^{(0)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Iteration 1. We have

$\mathbf{x}^{(0)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{y}^{(1)} = A\mathbf{x}^{(0)} = \begin{bmatrix} -5.0000 \\ -3.0000 \end{bmatrix}$, $c_1 = -5$, $\mathbf{x}^{(1)} = \begin{bmatrix} 1.6666 \\ 1.0000 \end{bmatrix}$.

Iteration 2.

$$\mathbf{y}^{(2)} = A\mathbf{x}^{(1)} = \begin{bmatrix} 1.6666 \\ 0.3333 \end{bmatrix},\ c_2 = 1.6666,\ \mathbf{x}^{(2)} = \begin{bmatrix} 1.0000 \\ 0.2000 \end{bmatrix}.$$

Iteration 3.

$$\mathbf{y}^{(3)} = A\mathbf{x}^{(2)} = \begin{bmatrix} 1.6666 \\ 1.4000 \end{bmatrix},\ c_3 = 1.6666,\ \mathbf{x}^{(3)} = \begin{bmatrix} 1.0000 \\ 0.4600 \end{bmatrix}.$$

Iteration 4.

$$\mathbf{y}^{(4)} = A\mathbf{x}^{(3)} = \begin{bmatrix} 1.6666 \\ 1.6000 \end{bmatrix},\ c_4 = 1.6666,\ \mathbf{x}^{(4)} = \begin{bmatrix} 1.0000 \\ 0.3600 \end{bmatrix}.$$

Iteration 5.

$$\mathbf{y}^{(5)} = A\mathbf{x}^{(4)} = \begin{bmatrix} 2.2000 \\ 0.92000 \end{bmatrix},\ c_5 = 2.2000,\ \mathbf{x}^{(5)} = \begin{bmatrix} 1.0000 \\ 0.4181 \end{bmatrix}.$$

Iteration 6.

$$\mathbf{y}^{(6)} = A\mathbf{x}^{(5)} = \begin{bmatrix} 1.9090 \\ 0.74545 \end{bmatrix},\ c_5 = 1.9090,\ \mathbf{x}^{(6)} = \begin{bmatrix} 1.0000 \\ 0.39047 \end{bmatrix}.$$

Continuing this way the 10th iterate is

$$\mathbf{y}^{(10)} = A\mathbf{x}^{(9)} = \begin{bmatrix} 1.99415 \\ 0.79649 \end{bmatrix},\ c_{10} = 1.9915,\ \mathbf{x}^{(10)} = \begin{bmatrix} 1.0000 \\ 0.399417 \end{bmatrix}.$$

Clearly it means the dominant eigenvalue is approaching to 2 and the corresponding dominant eigenvector is approaching to $\begin{bmatrix} 1.0 \\ 0.40 \end{bmatrix}$.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
from numpy import argmax,argmin
A=matrix([[4,-5],[2,-3]])
x0=vector([0.0,1.0]) # Initial guess of eigenvector
maxit=20 # Maximum number of iterates
dig=8 # number of decimal places to be shown is dig-1
tol=0.0001
 # Tolerance limit for difference of two consecutive eigenvectors
err=1 # Initialization of tolerance
i=0
while(i<=maxit and err>=tol):
    y0=A*x0
    ymod=y0.apply_map(abs)
    imax=argmax(ymod)
    c1=y0[imax]
    x1=y0/c1
    err=norm(x0-x1)
    i=i+1
```

```
    x0=x1
    print "Iteration Number:", i-1
    print  "y"+str(i-1)+"=",y0.n(digits=dig),  "c"+str(i-1)+"=",  c1.n(digits=dig),
"x"+str(i)+"=",x0.n(digits=dig)
    print "n"
```

```
Iteration Number: 0
y0= (-5.0000000, -3.0000000) c0= -5.0000000 x1= (1.0000000, 0.60000000)
n
Iteration Number: 1
y1= (1.0000000, 0.20000000) c1= 1.0000000 x2= (1.0000000, 0.20000000)
n
Iteration Number: 2
y2= (3.0000000, 1.4000000) c2= 3.0000000 x3= (1.0000000, 0.46666667)
n
Iteration Number: 3
y3= (1.6666667, 0.60000000) c3= 1.6666667 x4= (1.0000000, 0.36000000)
n
Iteration Number: 4
y4= (2.2000000, 0.92000000) c4= 2.2000000 x5= (1.0000000, 0.41818182)
n
Iteration Number: 5
y5= (1.9090909, 0.74545455) c5= 1.9090909 x6= (1.0000000, 0.39047619)
n
Iteration Number: 6
y6= (2.0476190, 0.82857143) c6= 2.0476190 x7= (1.0000000, 0.40465116)
n
Iteration Number: 7
y7= (1.9767442, 0.78604651) c7= 1.9767442 x8= (1.0000000, 0.39764706)
n
Iteration Number: 8
y8= (2.0117647, 0.80705882) c8= 2.0117647 x9= (1.0000000, 0.40116959)
n
Iteration Number: 9
y9= (1.9941520, 0.79649123) c9= 1.9941520 x10= (1.0000000, 0.39941349)
n
Iteration Number: 10
y10= (2.0029326, 0.80175953) c10= 2.0029326 x11= (1.0000000, 0.40029283)
n
Iteration Number: 11
```

```
y11= (1.9985359, 0.79912152) c11= 1.9985359 x12= (1.0000000, 0.39985348)
n
Iteration Number: 12
y12= (2.0007326, 0.80043956) c12= 2.0007326 x13= (1.0000000, 0.40007323)
n
Iteration Number: 13
y13= (1.9996338, 0.79978030) c13= 1.9996338 x14= (1.0000000, 0.39996338)
n
Iteration Number: 14
y14= (2.0001831, 0.80010987) c14= 2.0001831 x15= (1.0000000, 0.40001831)
n
```
■

Example 2

Let us find the dominant eigenvalue and the corresponding eigenvector of the matrix $A=\begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$, starting with $\mathbf{x}^{(0)}=\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Iteration 1. We have

$$\mathbf{x}^{(0)}=(1,1,1),\ \mathbf{y}^{(1)}=A\mathbf{x}^{(0)}=\begin{bmatrix} 1.00000 \\ 1.00000 \\ 4.0000 \end{bmatrix},\ c_1=4.0,\ \mathbf{x}^{(1)}=\begin{bmatrix} 0.25000 \\ 0.25000 \\ 1.00000 \end{bmatrix}.$$

Iteration 2. $\mathbf{y}^{(2)}=A\mathbf{x}^{(1)}=\begin{bmatrix} 2.25000 \\ 2.25000 \\ 4.0000 \end{bmatrix},\ c_2=4.0,\ \mathbf{x}^{(2)}=\frac{1}{c_2}\mathbf{y}^{(2)}=\begin{bmatrix} 0.62500 \\ 0.62500 \\ 1.00000 \end{bmatrix}.$

Iteration 3. $\mathbf{y}^{(3)}=A\mathbf{x}^{(2)}=\begin{bmatrix} 1.75000 \\ 1.75000 \\ 4.0000 \end{bmatrix},\ c_3=4.0,\ \mathbf{x}^{(3)}=\frac{1}{c_3}\mathbf{y}^{(3)}=\begin{bmatrix} 0.43750 \\ 0.43750 \\ 1.00000 \end{bmatrix}.$

Iteration 4. $\mathbf{y}^{(4)}=A\mathbf{x}^{(3)}=\begin{bmatrix} 2.12500 \\ 2.12500 \\ 4.0000 \end{bmatrix},\ c_4=4.0,\ \mathbf{x}^{(4)}=\frac{1}{c_4}\mathbf{y}^{(4)}=\begin{bmatrix} 0.53125 \\ 0.53125 \\ 1.00000 \end{bmatrix}.$

Continuing this, the 10th iterate is given by

$$\mathbf{y}^{(10)}=A\mathbf{x}^{(9)}=\begin{bmatrix} 2.00195 \\ 2.00195 \\ 4.0000 \end{bmatrix},\ c_{10}=4.0,\ \mathbf{x}^{(10)}=\frac{1}{c_{10}}\mathbf{y}^{(10)}=\begin{bmatrix} 0.50048 \\ 0.50048 \\ 1.00000 \end{bmatrix}.$$

Clearly it means the dominant eigenvalue is approaching to 4 and the corresponding dominant eigenvector is approaching to $\begin{bmatrix} 0.5 \\ 0.5 \\ 1.0 \end{bmatrix}$.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
from numpy import argmax, argmin
A=matrix([[1,-3,3],[3, -5, 3],[6,-6,4]])
x0=vector([1.0,1.0,1.0]) ## Initial guess
```

```
maxit=20 # Maximum number of iterates
dig=8 # number of decimal places to be shown is dig-1
tol=0.00001
# Tolerance limit for difference of two consecutive eigenvectors
err=1 # Initialization of tolerance
i=0
while(i<=n and err>=tol):
    y0=A*x0
    ymod=y0.apply_map(abs)
    imax=argmax(ymod)
    c1=y0[imax]
    x1=y0/c1
    err=norm(x0-x1)
    i=i+1
    x0=x1
    print "Iteration Number:", i-1
    print "y"+str(i-1)+"=",y0.n(digits=dig), "   c"+str(i-1)+"=", c1.n(digits=dig)
    print "x"+str(i)+"=",x0.n(digits=dig)
    print "n"
```

[Remark]

The rate convergence of the power method is determined by the ration $\left|\dfrac{\lambda_2}{\lambda_1}\right|$. Smaller is the ratio better is the convergence.

Example **3** Consider matrices $A=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}$, $B=\begin{bmatrix}1.7 & -0.4\\ 0.15 & 2.2\end{bmatrix}$, $C=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}$.

Starting with arbitrary vector $\mathbf{x}^{(0)}$ observe that for A convergence is obtained in fewer iterates, for B the convergence requires many more iterates where as for C there is no convergence.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
Mat=['A','B','C']
from numpy import argmax,argmin
@interact
def  _QRMethod(A1=input_box(default='[[1,2],[3,4]]',  type  =  str,  label  =
```

```
'A'),B1=input_box(default='[[1.7,-0.4],[0.15,2.2]]', type = str, label = 'B'),C1=input_box(default='[[1,2],[-3,4]]', type = str, label = 'C'),example=selector(Mat,buttons=True,label='Choose the Matrix'),maxit=slider(1, 500, 1, default=100, label="Maximum no. of iterations"),tol=input_box(label="Tolerance",default=0.001),v= input_box([0.1,1.0])):
    if(example=='A'):
        A1=sage_eval(A1)
        A=matrix(A1)
    elif(example=='B'):
        B1=sage_eval(B1)
        A=matrix(B1)
    elif(example=='C'):
        C1=sage_eval(C1)
        A=matrix(C1)
    x0=vector(v)
    html('A=%s,~~ x_0=%s'%(latex(A),latex(x0)))
    #html('x_0=%s'%latex(x0))
    #x0=vector([0.0,1.0])
    i=0
    err=1
    while(i<=maxit and err>=tol):
        y0=A*x0
        ymod=y0.apply_map(abs)
        imax=argmax(ymod)
        c1=y0[imax]
        x1=y0/c1
        err=norm(x0-x1)
        print "Iteration Number:", i+1
        html('y_i=%s,~~ c_i=%s~~ x_i=%s'%(latex(y0),latex(c1),latex(x0)))
        i=i+1
        x0=x1
    if(i==maxit+1):
        print 'Convergence is not achieved'
    else:
        print 'The number iteration required for tolerance=',tol,'is:',i
```

[Remark] *nonzero smallest eigenvalue*

To find the non zero smallest eigenvalue of a matrix A, we can find the dominant eigenvalues of A^{-1}.

[Remark] *shifted power method*

Note that if λ is an eigenvalue of A then $\lambda - \sigma$ is an eigenvalue of $A - \sigma I$. The rate of convergence of a power method can be significantly improved by using a shifted matrix $A - \sigma I$ rather than A in the power method. This method is called the **shifted power method**.

Inverse Power Method

In case, a reasonably "good approximation" of an eigenvalue is known, then we can use the "inverse power method" to find an eigenvalue and the corresponding eigenvector.

Let σ be an approximation to an eigenvalue λ_1 such that $|\lambda_1 - \sigma| \ll |\lambda_i - \sigma|$ for all $i = 2, 3,$ ... That is, σ is much closer to λ_1 than to the other eigenvalues. Then we have the following algorithm

Inverse Power Method Algorithm

[Step 1] Select an initial estimate σ sufficiently close to λ_1.

[Step 2] Select the a vector $\mathbf{x}^{(0)}$ whose largest entry is 1.

[Step 3] Set $k = 0$.

[Step 4] Solve $(A - \sigma I)\mathbf{y}^{(k)} = \mathbf{x}^{(k)}$ for $\mathbf{y}^{(k)}$.

[Step 5] Define c_k to be largest component in absolute value in the vector $\mathbf{x}^{(k)}$.

[Step 6] Find $d_k = \sigma + \dfrac{1}{c_k}$.

[Step 7] Define $\mathbf{x}^{(k+1)} = \dfrac{1}{c_k}\mathbf{y}^{(k)}$.

[Step 8] Check if the convergence criteria is met. Otherwise

[Step 9] Set $k = k + 1$ and go the the step 4.

In the above algorithm d_k converges to λ_1 and $\mathbf{x}^{(k)}$ converges to the corresponding eigenvector.

The inverse power method is also know as **inverse iteration with shift method**.

[Remark]

Note that the inverse power method is nothing but the power method applied to the matrix $(A-\sigma I)^{-1}$. This is why the name.

Example 4

Consider the matrix $A=\begin{bmatrix} 10 & -8 & -4 \\ -8 & 13 & 5 \\ -4 & 4 & 4 \end{bmatrix}$. Suppose $\sigma = 1.8$ is an estimate of an eigenvalue of A. Apply the inverse power method to approximate an eigenvalue of A starting with $\mathrm{x}^{(0)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Iteration 1. $\mathrm{y}^{(0)} = (A-\sigma I)^{-1}\mathrm{x}^{(0)} = \begin{bmatrix} 3.02739 \\ -3.80292 \\ 13.48630 \end{bmatrix}$;

$$c_0 = 13.48630,\ d_0 = 1.97414,\ \mathrm{x}^{(1)} = \frac{1}{c_0}\mathrm{y}^{(0)} = \begin{bmatrix} 0.22448 \\ -0.28198 \\ 1.0000 \end{bmatrix}.$$

Iteration 2. $\mathrm{y}^{(1)} = (A-\sigma I)^{-1}\mathrm{x}^{(1)} = \begin{bmatrix} 1.74935 \\ -3.38069 \\ 10.24772 \end{bmatrix}$;

$$c_1 = 10.24772,\ d_1 = 1.99758,\ \mathrm{x}^{(2)} = \frac{1}{c_1}\mathrm{y}^{(1)} = \begin{bmatrix} 0.17071 \\ -0.32989 \\ 1.0000 \end{bmatrix}.$$

Iteration 3. $\mathrm{y}^{(2)} = (A-\sigma I)^{-1}\mathrm{x}^{(2)} = \begin{bmatrix} 1.67240 \\ -3.336698 \\ 10.01733 \end{bmatrix}$;

$$c_2 = 10.01733,\ d_2 = 1.99983,\ \mathrm{x}^{(3)} = \frac{1}{c_2}\mathrm{y}^{(2)} = \begin{bmatrix} 0.16695 \\ -0.33309 \\ 1.000 \end{bmatrix}.$$

Continuing this way, we have Iteration 10.

$$\mathrm{y}^{(9)} = (A-\sigma I)^{-1}\mathrm{x}^{(9)} = \begin{bmatrix} 1.66666 \\ -3.333333 \\ 10.0000 \end{bmatrix};$$

$$c_9 = 10.24772,\ d_9 = 1.999999,\ \mathrm{x}^{(10)} = \begin{bmatrix} 0.166666 \\ -0.333333 \\ 1.000 \end{bmatrix}.$$

This mean an approximate eigenvalue is 2 and the corresponding eigenvector converges to $\begin{bmatrix} 1/6 \\ 1/3 \\ 1 \end{bmatrix}$.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
from numpy import argmax, argmin
A=matrix([[10,-8,-4],[-8,13,5],[-4,4,4]])
Id=identity_matrix(3)
x0=vector([1.0,1.0,1.0]) ## Initial guess
maxit=20 # Maximum number of iterates
```

```
dig=8 # number of decimal places to be shown is dig-1
tol=0.00001
# Tolerance limit for difference of two consecutive eigenvectors
err=1 # Initialization of tolerance
sig=1.9 # Initial Shifting number
i=0
while(i<=n and err>=tol):
    y0=(A-sig*Id).inverse()*x0
    ymod=y0.apply_map(abs)
    imax=argmax(ymod)
    c1=y0[imax]
    d1=sig+1/c1
    x1=y0/c1
    print "Iteration Number:", i+1
    print "y"+str(i)+"=",y0.n(digits=dig), "d"+str(i)+"=", d1.n(digits=dig)
    print "x"+str(i+1)+"=",x0.n(digits=dig)
    print "n"
    i=i+1
    x0=x1
```

[Remark]

The advantage of the inverse power method with shift is that it can be adopted to find any eigenvalue of a matrix, instead of the extreme ones. However, in order to compute a particular eigenvalue, we need to have an initial approximation that of that eigenvalue.

Rayleigh Quotient

The Rayleigh quotient of a non zero vectors $\mathbf{x}$ with respect of a matrix is defines as

$$r(\mathbf{x}) = \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$$

If $\mathbf{x}$ is an eigenvector with respect to the eigenvalue λ, then $r(\mathbf{x}) = \lambda$.

In general, for an arbitrary $\mathbf{x}$, $r(\mathbf{x})$ is value that minimizes the function

$f(\lambda) = \| A\mathbf{x} - \lambda\mathbf{x} \|^2$ over real number λ.

The inverse power method can be significantly improved if we drop the restriction that the shift value σ remains constant in all the iterates.

Each iteration in the inverse power method gives and approximation of eigenvector, given an estimation of eigenvalue. On the other hand, the Rayleigh quotient gives an approximate eigenvalue, given as estimate of an eigenvector. Combining the two concepts together, we get a new variation in the inverse power algorithm in which the shift value σ is updated in each iterate and it becomes the Rayleigh quotient of the eigenvector estimates. This method is called the **Rayleigh quotient iteration method (RQI)**.

Rayleigh Quotient Iteration Algorithm

[Step 1] Select the a vector $\mathbf{x}^{(0)}$ with $\| \mathbf{x}^{(0)} \| = 1$.

[Step 2] Define $\lambda^{(0)} = r(\mathbf{x}^{(0)}) = (\mathbf{x}^{(0)})^T A\mathbf{x}^{(0)}$.

[Step 3] Set $k = 1$.

[Step 4] Define $\lambda^{(k-1)} = r(\mathbf{x}^{(k-1)}) = (\mathbf{x}^{(k-1)})^T A\mathbf{x}^{(k-1)}$.

[Step 5] Solve $(A - \lambda^{(k-1)} I)\mathbf{y}^{(k)} = \mathbf{x}^{(k-1)}$ for $\mathbf{y}^{(k)}$.

[Step 6] Define $\mathbf{x}^{(k)} = \dfrac{\mathbf{y}^{(k)}}{\| \mathbf{y}^{(k)} \|}$.

[Step 7] Check if the convergence criteria is met. Otherwise

[Step 8] Set $k = k+1$ and go the the step 4.

One of the main advantage of the RQI is that it converges much faster than power method and inverse power method. However, a very significant disadvantage of RQI is that its convergence is not always guaranteed except when the matrix is symmetric.

Example 5 Apply the Rayleigh quotient iteration method to find an eigenvalue of the matrix $A = \begin{bmatrix} 10 & -8 & -4 \\ -8 & 13 & 5 \\ -4 & 5 & 4 \end{bmatrix}$ starting with initial approximate vector $\mathbf{x}^{(0)} = \begin{bmatrix} 1.5 \\ -2.5 \\ 5 \end{bmatrix}$.

Iteration 1. $\mathbf{x}^{(0)} = \begin{bmatrix} 1.5 \\ -2.5 \\ 5 \end{bmatrix}$ and normalize $\mathbf{x}^{(0)}$ to get

$$\mathbf{x}^{(0)} = \frac{\mathbf{x}^{(0)}}{\| \mathbf{x}^{(0)} \|} = \begin{bmatrix} 0.259160527674408 \\ -0.431934212790680 \\ 0.863868425581360 \end{bmatrix}.$$

Now $\lambda^{(0)} = r(\mathbf{x}^{(0)}) = (\mathbf{x}^{(0)})^T A \mathbf{x}^{(0)} = 2.35074626865672$.

Solving $(A - \lambda^{(0)} I)\mathbf{y}^{(1)} = \mathbf{x}^{(0)}$ for $\mathbf{y}^{(1)}$ we get, $\mathbf{y}^{(1)} = \begin{bmatrix} -0.253865948919549 \\ 0.444271881412964 \\ -1.43880515900300 \end{bmatrix}$.

Hence $\mathbf{x}^{(1)} = \dfrac{\mathbf{y}^{(1)}}{\| \mathbf{y}^{(1)} \|} = \begin{bmatrix} -0.166242289108663 \\ 0.290928243299417 \\ -0.942191200639566 \end{bmatrix}$.

Iteration 2. $\lambda^{(1)} = r(\mathbf{x}^{(1)}) = (\mathbf{x}^{(1)})^T A \mathbf{x}^{(1)} = 1.7072489$.

Solving $(A - \lambda^{(1)} I)\mathbf{y}^{(2)} = \mathbf{x}^{(1)}$ for $\mathbf{y}^{(2)}$ we get, $\mathbf{y}^{(2)} = \begin{bmatrix} 201.04056 \\ -367.87171 \\ 1152.5798 \end{bmatrix}$.

Hence $\mathbf{x}^{(2)} = \dfrac{\mathbf{y}^{(2)}}{\| \mathbf{y}^{(2)} \|} = \begin{bmatrix} 0.16392031 \\ -0.29994765 \\ 0.93976675 \end{bmatrix}$.

Iteration 3. $\lambda^{(2)} = r(\mathbf{x}^{(2)}) = (\mathbf{x}^{(2)})^T A \mathbf{x}^{(2)} = 1.7064336$.

Solving $(A - \lambda^{(1)} I)\mathbf{y}^{(3)} = \mathbf{x}^{(2)}$ for $\mathbf{y}^{(3)}$ and then we have

$\mathbf{x}^{(3)} = \dfrac{\mathbf{y}^{(3)}}{\| \mathbf{y}^{(3)} \|} = \begin{bmatrix} -0.163923 \\ 0.299945 \\ -0.939767 \end{bmatrix}$.

Iteration 4. $\lambda^{(3)} = 1.70643$ and $\mathbf{x}^{(4)} = \begin{bmatrix} -0.163923 \\ 0.299945 \\ -0.939767 \end{bmatrix}$.

Clearly, in 4 iterates we are getting reasonably accurate eigenvalue $\lambda = 1.70643$ and the corresponding eigenvector $\begin{bmatrix} -0.163923 \\ 0.299945 \\ -0.939767 \end{bmatrix}$.

Other Eigenvalues

We can use different initial vectors $\mathbf{x}^{(0)}$ to get a different eigenvalues and the corresponding eigenvectors.

For example if we use $\mathbf{x}^{(0)} = \begin{bmatrix} 1.0 \\ 0 \\ 0 \end{bmatrix}$, then after 4 iterates we have

$$\lambda^{(3)} = 3.36758, \ \mathbf{x}^{(4)} = \begin{bmatrix} -0.783163 \\ -0.618827 \\ -0.0609041 \end{bmatrix}.$$

If we use $\mathbf{x}^{(0)} = \begin{bmatrix} 1.0 \\ -1 \\ -1 \end{bmatrix}$, then after 4 iterates we have

$$\lambda^{(3)} = 21.9259, \quad \mathbf{x}^{(4)} = \begin{bmatrix} -0.599821 \\ 0.726007 \\ 0.336346 \end{bmatrix}.$$

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix([[10,-8,-4],[-8,13,5],[-4,5,4]])
Id=identity_matrix(3)
x0=vector([1.5,-2.5,5])
#x0=vector([1.0,0.0,0.0])
#x0=vector([1.0,-1,-1])
x0=x0/norm(x0)
maxit=20 # Maximum number of iterates
dig=8 # number of decimal places to be shown is dig-1
tol=0.00001
# Tolerance limit for difference of two consecutive eigenvectors
err=1 # Initialization of tolerance
i=0
while(i<=n and err>=tol):
    lam0=x0.dot_product(A*x0)
    y0=(A-lam0*Id).inverse()*x0
    x1=y0/norm(y0)
    print "Iteration Number:", i+1
    print "y"+str(i)+"=",y0.n(digits=dig), "lambda"+str(i)+"=", lam0.n(digits=dig)
    print "x"+str(i+1)+"=",x0.n(digits=dig)
    print "n"
    i=i+1
    x0=x1
```

QR Method

The QR method for finding eigenvalues and eigenvectors is a simultaneous iteration method that allows us to find all eigenvalues and eigenvectors of a real, symmetric full rank matrix at once.

The algorithm is simple:

- We start with $A^{(0)} = A$.
- Set $k=1$ and find the QR factorization $Q^{(k)}R^{(k)} = A^{(k-1)}$.
- Let $A^{(k)} = R^{(k)}Q^{(k)}$.

The sequence $A^{(k)}$ has the following properties: for each k, $A^{(k)}$ is orthogonally equivalent to $A^{(k-1)}$ and hence is orthogonally equivalent to the original matrix A.

$A^{(1)} = R^{(0)}Q^{(0)} = (Q^{(0)})^T A^{(0)} Q^{(0)}$ since $R^{(0)} = (Q^{(0)})^T A^{(0)}$.

Similarly,

$$A^{(2)} = R^{(1)}Q^{(1)} = (Q^{(1)})^T A^{(1)} Q^{(1)}.$$

It can be shown that the sequence $A^{(k)}$ converges (under certain conditions) to an upper triangular matrix or quasi-triangular matrix. In particular, the diagonal entries of $A^{(k)}$ are eigenvalues.

If we define

$$\hat{Q}^{(k)} := Q^{(1)}Q^{(2)}\cdots Q^{(k)}$$

Then the columns of $\hat{Q}^{(k)}$ converges to unit eigenvectors of A.

Similarly we can define

$$\hat{R}^{(k)} := R^{(k)}R^{(k-1)}\cdots R^{(2)}R^{(1)}.$$

[Remark] (i) $A^{(k)} = \hat{Q}^{(k)}\hat{R}^{(k)}$ and (ii) $A^{(k)} = \left(\hat{Q}^{(k)}\right)^T A \hat{Q}^{(k)}$.

Example 6

Consider $A = \begin{bmatrix} 10.0 & 3.0 & 4.0 \\ 3.0 & 5.0 & 1.0 \\ 4.0 & 2.0 & 3.0 \end{bmatrix}$. The actual eigenvalues of A are 13.1804689044, 3.56330346867, 1.25622762694.

Now we apply the QR-method.

Iteration 1. $A^{(0)} = Q^{(0)}R^{(0)}$. We have

$$Q^{(0)} = \begin{bmatrix} -0.894427191 & 0.314676219522 & -0.31777173705 \\ -0.2683281573 & -0.946058827723 & -0.181583849743 \\ -0.3577708764 & -0.0771464280117 & 0.930617229931 \end{bmatrix}.$$

$$R^{(0)} = \begin{bmatrix} -11.1803398875 & -4.7404641123 & -4.9193495505 \\ 0.0 & -3.94055833607 & 0.0812067663282 \\ 0.0 & 0.0 & 1.33918089185 \end{bmatrix}.$$

$$A^{(1)} = R^{(0)}Q^{(0)} = \begin{bmatrix} 13.032 & 1.34608107814 & -0.164443701815 \\ 1.02830934109 & 3.72173518805 & 0.791114168732 \\ -0.479119921336 & -0.103313022268 & 1.24626481195 \end{bmatrix}.$$

Iteration 2. $A^{(1)} = Q^{(1)}R^{(1)}$. We have

$$Q^{(1)} = \begin{bmatrix} -0.996232453991 & 0.0790956260904 & 0.0355637392449 \\ -0.0786092033716 & -0.996794511229 & 0.0148760051455 \\ 0.0366263670188 & 0.0120243219007 & 0.999256686203 \end{bmatrix}.$$

$$R^{(1)} = \begin{bmatrix} -13.0812843406 & -1.63735627471 & 0.147281450427 \\ 0.0 & -3.60457835109 & -0.786599549427 \\ -0.0 & -0.0 & 1.25125883164 \end{bmatrix}.$$

$$A^{(2)} = R^{(1)}Q^{(1)} = \begin{bmatrix} 13.1661056568 & 0.599206332142 & -0.342404731552 \\ 0.254542748875 & 3.58356558947 & -0.839636585227 \\ 0.045829065203 & 0.0150455389727 & 1.25032875368 \end{bmatrix}.$$

Iteration 3. $A^{(2)} = Q^{(2)}R^{(2)}$. We have

$$Q^{(2)} = \begin{bmatrix} -0.999807111701 & 0.0193418421557 & -0.00341065019922 \\ -0.0193294552839 & -0.999806585683 & -0.00362813765143 \\ -0.00348016539642 & -0.00356151181562 & 0.999987601964 \end{bmatrix}.$$

$$R^{(2)} = \begin{bmatrix} -13.1686457345 & -0.668411484034 & 0.354217052652 \\ 0.0 & -3.57133630715 & 0.82839838859 \\ 0.0 & 0.0 & 1.25452739194 \end{bmatrix}.$$

$$A^{(3)} = R^{(2)}Q^{(2)} = \begin{bmatrix} 13.1777929528 & 0.412314788263 & 0.401551394126 \\ 0.0661490220464 & 3.56769520893 & 0.841345417799 \\ -0.00436596281828 & -0.0044680141294 & 1.25451183826 \end{bmatrix}.$$

Continuing this iterations, in 20th iterate we have $A^{(19)} = Q^{(19)}R^{(19)}$.
We have

$$Q^{(19)} = \begin{bmatrix} -1.0 & 3.99321544727\times10^{-12} & 1.50791309447\times10^{-20} \\ -3.99321544727\times10^{-12} & -1.0 & 6.8298793812\times10^{-11} \\ 1.53518627432\times10^{-20} & 6.8298793812\times10^{-11} & 1.0 \end{bmatrix}.$$

$$R^{(19)} = \begin{bmatrix} -13.1804689044 & -0.345707260572 & -0.41234692193 \\ 0.0 & -3.56330346876 & -0.842885819939 \\ -0.0 & -0.0 & 1.25622762691 \end{bmatrix}.$$

$$A^{(20)} = R^{(19)}Q^{(19)}$$
$$= \begin{bmatrix} 13.1804689044 & 0.345707260491 & -0.412346921953 \\ 1.42290384418\times10^{-11} & 3.5633034687 & -0.842885820182 \\ 1.92854341025\times10^{-20} & 8.57988316715\times10^{-11} & 1.25622762691 \end{bmatrix}.$$

Clearly, diagonal entries of $A^{(20)}$ are close to actual eigenvalues of A.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080

```
A=matrix(RDF, [[10,3,4],[3,5,1],[4,2,3]])
print "The actual eivenvalues of A are"
ev=A.eigenvalues()
show(ev)
n=10
for i in range(n):
    Q1,R1=A.QR()
    print "Iteration Number", i
    print "The matrix Q"+str(i), "is"

    show(Q1)
    print "The matrix R"+str(i), "is"
    show(R1)
    A1=R1*Q1
    print "The matrix A"+str(i), "is"
    show(A1)
    A=A1
```

[Remark]

QR method mentioned above usually is very expensive. This is why usually, symmetric matrices are first converted to tridiagonal matrix and then we apply QR method. For non-symmetric matrix, we convert it to an upper Hessenberg matric and then apply QR method.

http://www.prenhall.com/bretscher1e/html/proj10.html

5.3 *Linear Model

● Reference video: http://youtu.be/CLxjkZuNJXw

● Practice site: http://matrix.skku.ac.kr/2012-LAwithSage/interact/
http://math1.skku.ac.kr/home/pub/1516/
http://matrix.skku.ac.kr/SOCW-Math-Modelling.htm

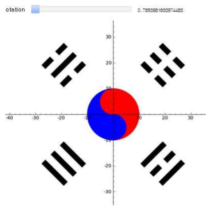
(1) Linear Algebra behind Google

http://www.rose-hulman.edu/~bryan/googleFinalVersionFixed.pdf
by Kurt Bryan and Tanya Leise

Google's success derives in large part from its PageRank algorithm, which ranks the importance of webpages according to an eigenvector of a weighted link matrix. Analysis of the PageRank formula provides a wonderful applied topic for a linear algebra course. Instructors may assign this article as a project to students, or spend one or two lectures presenting the material with assigned homework from the exercises. This material also complements the discussion of Markov chains in matrix algebra. Maple and Mathematica files supporting this material can be found in http://www.rose-hulman.edu/~bryan/googleFinalVersionFixed.pdf.

A newsletter article '"Linear Algebra and Google Search Engine" - Pagerank algorithm' in Korean also can be found in
http://matrix.skku.ac.kr/2012-e-Books/KMS-News-LA-Google-SGLee.pdf.

(2) Sage Matrix Calculator for Linear Algebra



In this section, we introduce a matrix calculator. By utilizing a free open source tool Sage, one can intuitively understand almost all concepts in linear algebra. Also, one can study with visualization and large scale computation. Moreover, one can easily change and expand the size of a matrix.

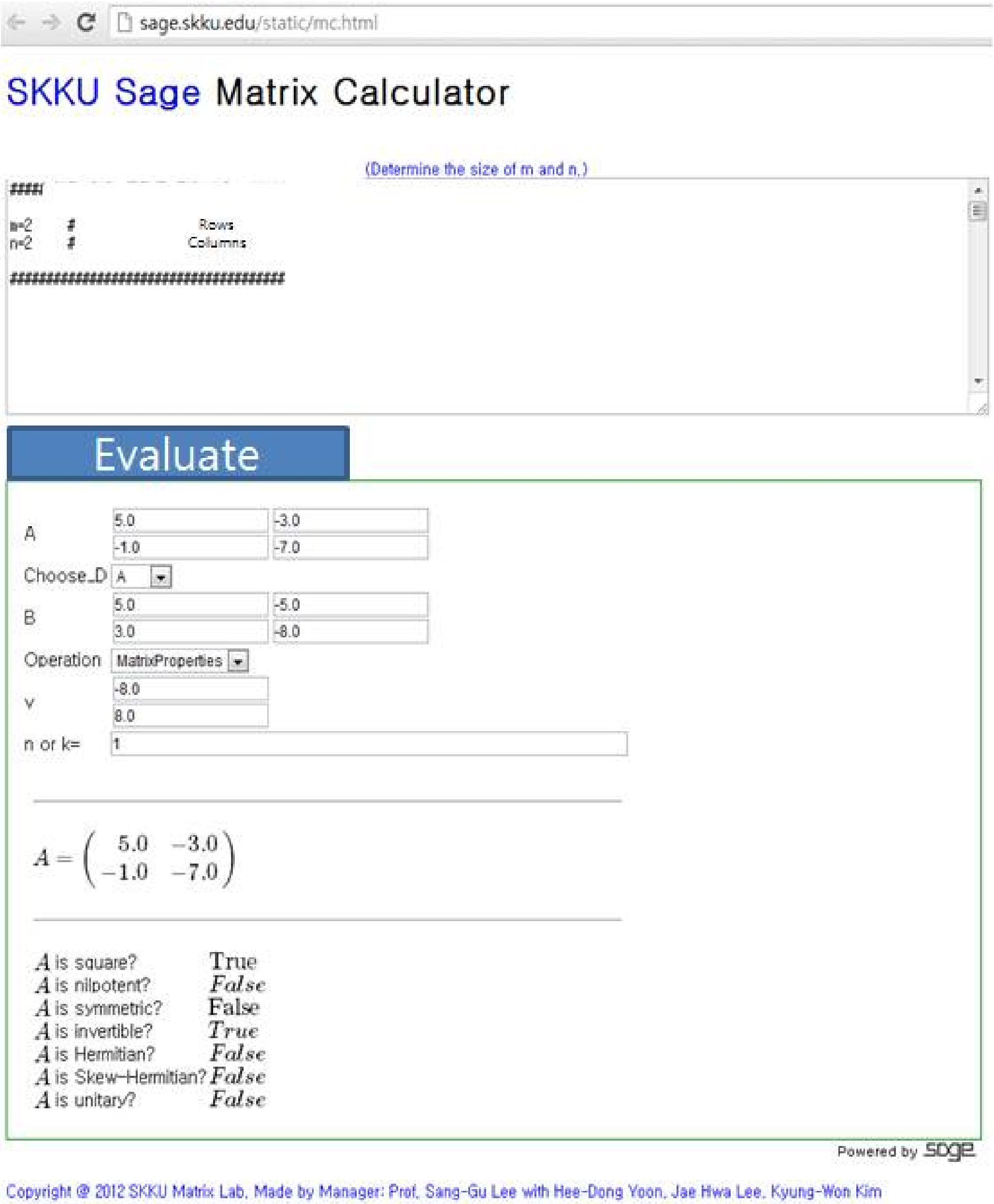
Sage Matrix Calculator

http://matrix.skku.ac.kr/2014-Album/MC.html

For over 20 years, the issue of using an adequate CAS tool in the teaching and learning of linear algebra has been raised constantly. A variety of CAS tools were introduced in many linear algebra textbooks; however, in Korea, due to some realistic problems, these tools have not been introduced in the class and the theoretical aspect of linear algebra has been the primary focus in Linear Algebra courses.

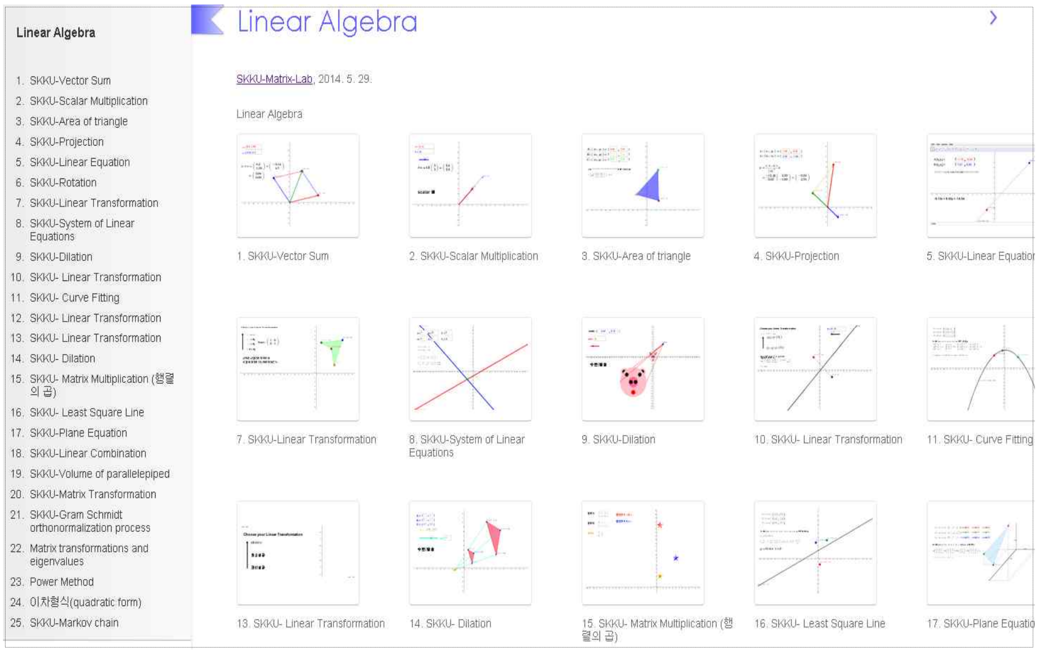
In this section, we suggest Sage as an alternative for CAS tools to overcome tthe problems mentioned above. As well, we introduce the extensive linear algebra content and a matrix calculator that was developed with Sage. Taking advantage of these novel tools, almost all concepts of linear algebra can be easily covered, and the size of matrices can be expanded without difficulty.

The Sage Matrix Calculator uses the Sage Cell server. As shown in the following picture, it can do not only basic operations, such as matrix addition, subtraction, multiplication, scalar multiplication, but also can find determinant, rank, trace, nullity, eigenvalues, characteristic equation, inverse matrix, adjoint matrix, transpose of matrix, and conjugate transpose of a matrix. Also, unlike most web-based open matrix calculators, it can perform LU, SVD, and QR-decomposition, which are quite essential to a well-rounded linear algebra education. By selecting the column size as 1, it can perform vector operations, such as inner product, cross product, and norm. As well, by using the column vectors of a matrix, it can perform Gram-Schmidt orthogonal process, and as a result, one can find the basis of a vector space generated by the matrix. As this matrix calculator can cover complex numbers, while many other matrix calculators can handle only real or rational numbers, it can solve almost all problems in linear algebra. In order to use the Sage matrix calculator, one needs only to connect to the given URL, or simply copy the codes from the given URL and paste



them to other Sage Cell server or a general Sage server's worksheet. Once it executed, decide the size of the matrix, enter the elements of the matrix, and then perform the desired matrix operations.

Visualization of Linear Algebra Concepts with GeoGebra



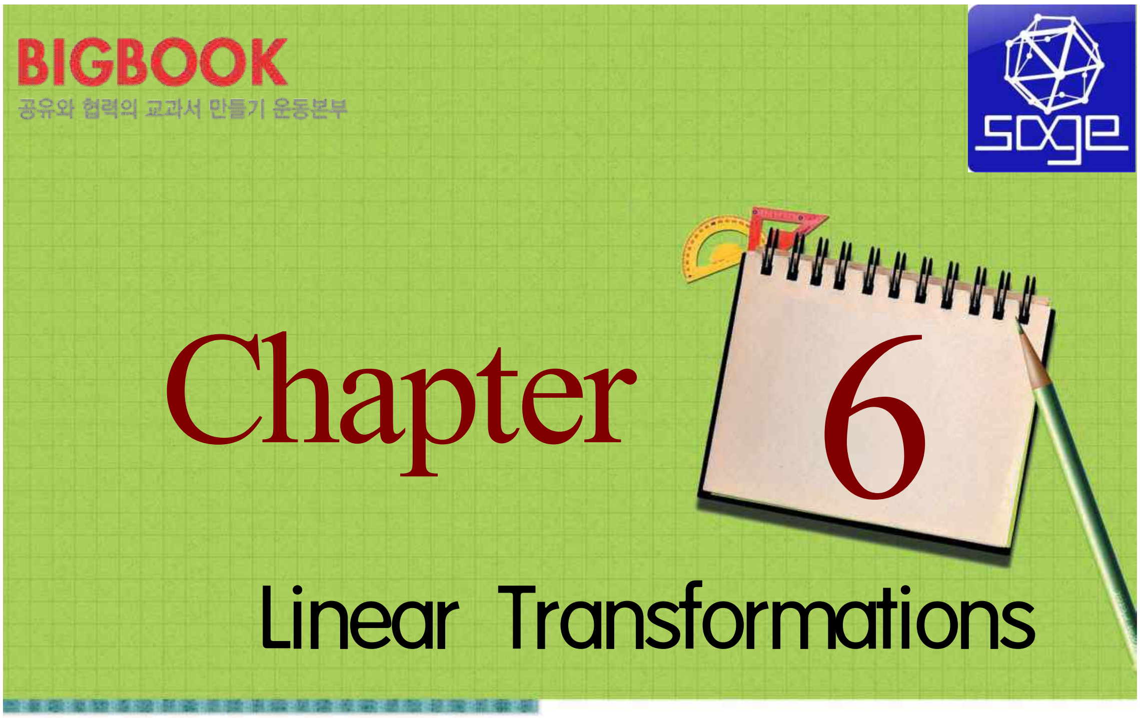
http://www.geogebratube.org/student/b121550

Vector addition	http://www.geogebratube.org/student/m9493
Sclar multiplication	http://www.geogebratube.org/student/m9494
L. S. of Equations	http://www.geogebratube.org/student/m9704
Matrix product	http://www.geogebratube.org/student/m12831
Areas	http://www.geogebratube.org/student/m9497
Equations	http://www.geogebratube.org/student/m9504
Curve Fitting	http://www.geogebratube.org/student/m9911
Linear Transformation	http://www.geogebratube.org/student/m9702
Projection	http://www.geogebratube.org/student/m9910
LT (Shear)	http://www.geogebratube.org/student/m9912
대칭변환	http://www.geogebratube.org/student/m9703
LT(similarity)	http://www.geogebratube.org/student/m9705
Triangles	http://www.geogebratube.org/student/11568
Projections	http://www.geogebratube.org/student/m9503
Least Square solution	http://www.geogebratube.org/student/m12933

http://matrix.skku.ac.kr/2012-Album/CLA-GeoGebra-Dynamic-Visual.htm

[References]

Sang-Gu LEE*, Kyung-Won KIM and Jae Hwa LEE, Sage matrix calculator and full Sage contents for linear algebra, *Korean J. Math*. 20 (2013), No. 4, pp. 503-521.



Chapter 6 Linear Transformations

So far, we have considered matrix mainly as a coefficient matrix from systems of linear equations. Now, we consider matrix as a function.

We have observed that the set of vectors and two operations reborn as an algebraic structure, namely a vector space. Matrix will be reborn as a linear transformation, which is a function that preserves the operations in a vector space. And linear transformations are used for noise filtering in signal processing and analysis in engineering processes.

We show a linear transformation from n-dimensional space $\mathbb{R}^n$ to m-dimensional space $\mathbb{R}^m$ can be expressed as a $m \times n$ matrix A. We shall also look at geometric meaning of linear transformations from $\mathbb{R}^2$ to $\mathbb{R}^2$ and applications in computer graphics.

6.1 Matrix as a Function (Transformation)

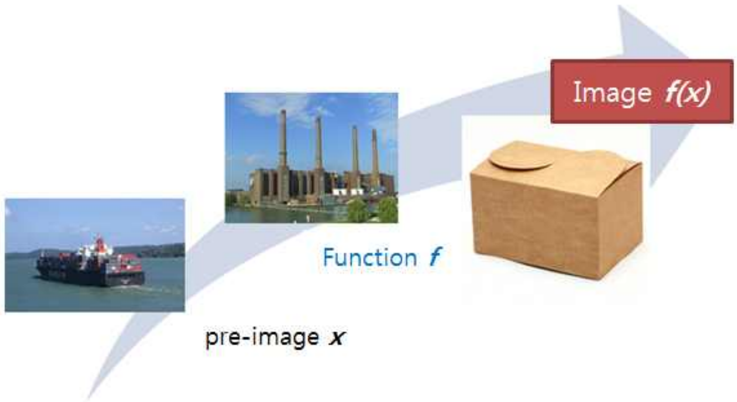
- Reference video: http://youtu.be/YF6-ENHfI6E, http://youtu.be/Yr23NRSpSoM
- Practice site: http://matrix.skku.ac.kr/knou-knowls/cla-week-8-Sec-6-1.html

Matrix can be considered as a special function with linearity property. Such a function play an important role in science and various areas in daily life, such as mathematics, physics, engineering control theory, image processing, sound signal, and computer graphics.

What is a Transformation?

Definition

A function, whose input and output are both vectors, is called a transformation. For a transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{w} = T(\mathbf{x})$ is called an image of $\mathbf{x}$ by T, and $\mathbf{x}$ is called a pre-image of $\mathbf{w}$.



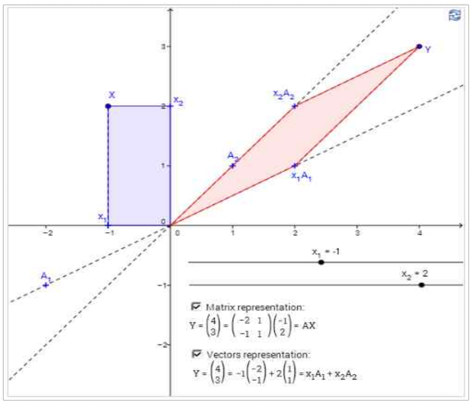
As a special case of transformations, $T_A(\mathbf{x}) = A\mathbf{x}$, for $m \times n$ matrix A and $\mathbf{x} \in \mathbb{R}^n$, $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is called a matrix transformation.

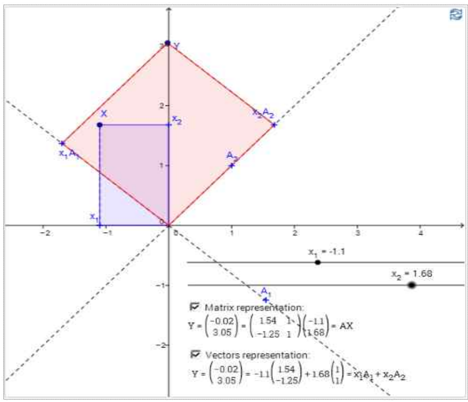
$$\mathbf{x}' = f(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} x_{n+1} \\ y_{n+1} \\ z_{n+1} \\ w_{n+1} \\ 1 \end{bmatrix} = \begin{bmatrix} a_{xx} & a_{yx} & a_{zx} & a_{wx} & a_x \\ a_{xy} & a_{yy} & a_{zy} & a_{wy} & a_y \\ a_{xz} & a_{yz} & a_{zz} & a_{wz} & a_z \\ a_{xw} & a_{yw} & a_{zw} & a_{ww} & a_w \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \\ z_n \\ w_n \\ 1 \end{bmatrix}$$

[Remark] Computer simulation

[Matrix transformation)

http://www.geogebratube.org/student/b73259#material/22419





Definition

If a transformation $T:\mathbb{R}^n \to \mathbb{R}^m$ from $\mathbb{R}^n$ to $\mathbb{R}^m$, satisfies the following two conditions for any vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and for any scalar $k \in \mathbb{R}$,

(1) $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ (2) $T(k\mathbf{u}) = kT(\mathbf{u})$

then T is called a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. Especially, a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$ itself, $T:\mathbb{R}^n \to \mathbb{R}^n$ is called a linear operator on $\mathbb{R}^n$.

Example 1

Show that T is a linear transformation if we define $T:\mathbb{R}^2 \to \mathbb{R}^3$, for any vector $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ in $\mathbb{R}^2$, as follows

$$T(\mathbf{x}) = \begin{bmatrix} x \\ y \\ x-y \end{bmatrix}$$

Solution

For any two vectors $\mathrm{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, $\mathrm{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ in $\mathbb{R}^2$ and for any scalar $k \in \mathbb{R}$,

$$(1)\ \ T(\mathrm{u}+\mathrm{v}) = T\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \end{bmatrix}\right)$$

$$= \begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ (u_1+v_1)-(u_2+v_2) \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_1-u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_1-v_2 \end{bmatrix}$$

$$= T(\mathrm{u}) + T(\mathrm{v}).$$

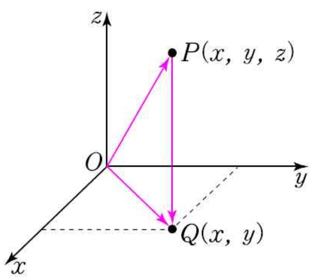
$$(2)\ \ T(k\mathrm{u}) = T\left(k\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} ku_1 \\ ku_2 \end{bmatrix}\right) = \begin{bmatrix} ku_1 \\ ku_2 \\ ku_1-ku_2 \end{bmatrix} = k\begin{bmatrix} u_1 \\ u_2 \\ u_1-u_2 \end{bmatrix} = kT(\mathrm{u}).$$

Therefore, by definition, T is a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$. ■

Example 2 Let $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T(x,y,z) = (x,y)$. Show that T is a linear transformation.

Solution

For any two vectors $\mathrm{v}_1 = (x_1, y_1, z_1)$ and $\mathrm{v}_2 = (x_2, y_2, z_2)$ in $\mathbb{R}^3$ and for any scalar $k \in \mathbb{R}$,



정사영

$$(1)\ \ T(\mathrm{v}_1+\mathrm{v}_2) = T(x_1+x_2,\ y_1+y_2,\ z_1+z_2)$$
$$= (x_1+x_2,\ y_1+y_2)$$
$$= (x_1, y_1) + (x_2, y_2) = T(\mathrm{v}_1) + T(\mathrm{v}_2).$$

$$(2)\ \ T(k\mathrm{v}) = T(kx,\ ky,\ kz)$$
$$= (kx,\ ky) = k(x,\ y) = k\,T(\mathrm{v})$$

Therefore, T is a linear transformation. ■

- This type of linear transformation is called orthogonal projection on xy-plane.

Example 3

If we define $T : \mathbb{R}^2 \to \mathbb{R}^2$ as follows, show that T is not a linear transformation.

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y+1 \end{bmatrix}.$$

Solution

For any two vectors, $v_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$, $v_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$ in $\mathbb{R}^2$,

$$T(v_1 + v_2) = T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 + 1 \end{bmatrix}.$$

However, $T(v_1) + T(v_2) = T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right) + T\left(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ y_1 + 1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 + 1 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 + 2 \end{bmatrix}$.

Hence $T(v_1 + v_2) \neq T(v_1) + T(v_2)$.

Therefore, we conclude that T is not a linear transformation. ■

[Remark] Special Linear Transformations

- zero transformation: For any $v \in \mathbb{R}^n$, if we define $T: \mathbb{R}^n \to \mathbb{R}^m$ as $T(v) = 0$, then T is a linear transformation. This is called a **zero transformation**.

- identity transformation: For any $v \in \mathbb{R}^n$, if we define $T: \mathbb{R}^n \to \mathbb{R}^m$ as $T(v) = v$, then T is a linear transformation. This is called an identity **transformation**.

- matrix transformation: For any $m \times n$ matrix A and for any vector $\mathbf{x}$ in $\mathbb{R}^n$, if we define $T_A(\mathbf{x}) = A\mathbf{x}$, then T_A is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. This is called a **matrix transformation**.

Example 4 Let $T: \mathbb{R}^3 \to \mathbb{R}^2$ is defined as follows. Show T is a linear transformation.

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Solution

For any two vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$, $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ in $\mathbb{R}^3$ and for any scalar $k \in \mathbb{R}$,

$$(1)\quad T(\mathrm{u}+\mathrm{v})=\begin{bmatrix}1&1&0\\0&1&-1\end{bmatrix}\begin{bmatrix}u_1+v_1\\u_2+v_2\\u_3+v_3\end{bmatrix}=\begin{bmatrix}1&1&0\\0&1&-1\end{bmatrix}\left(\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}+\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\right)$$

$$=\begin{bmatrix}1&1&0\\0&1&-1\end{bmatrix}\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}+\begin{bmatrix}1&1&0\\0&1&-1\end{bmatrix}\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}=T(\mathrm{u})+T(\mathrm{v})$$

$$(2)\quad T(k\mathrm{u})=\begin{bmatrix}1&1&0\\0&1&-1\end{bmatrix}\begin{bmatrix}ku_1\\ku_2\\ku_3\end{bmatrix}=k\begin{bmatrix}1&1&0\\0&1&-1\end{bmatrix}\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}=kT(\mathrm{u})$$

and hence, T is a linear transformation. ■

Example 5 A linear transformation T from Example 1 is $T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\begin{bmatrix}x\\y\\x-y\end{bmatrix}=\begin{bmatrix}1&0\\0&1\\1&-1\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}$,

T is a matrix transformation for a matrix $A=\begin{bmatrix}1&0\\0&1\\1&-1\end{bmatrix}$. ■

Theorem 6.1.1 [Properties of Linear Transformation]

If $T:\mathbb{R}^n\to\mathbb{R}^m$ is a linear transformation, then it satisfies the following conditions:

(1) $T(\mathbf{0})=\mathbf{0}$.
(2) $T(-\mathrm{v})=-T(\mathrm{v})$.
(3) $T(\mathrm{u}-\mathrm{v})=T(\mathrm{u})-T(\mathrm{v})$.

Proof (1) Since $\forall\, \mathrm{v}\in V$, $0\mathrm{v}=\mathbf{0}$, $T(\mathbf{0})=T(0\mathrm{v})=0\,T(\mathrm{v})=\mathbf{0}$.
(2) $T(-\mathrm{v})=T((-1)\mathrm{v})=(-1)\,T(\mathrm{v})=-T(\mathrm{v})$
(3) $T(\mathrm{u}-\mathrm{v})=T(\mathrm{u}+(-1)\mathrm{v})=T(\mathrm{u})+(-1)\,T(\mathrm{v})=T(\mathrm{u})-T(\mathrm{v})$ ■

Each linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be expressed as a matrix transformation.

- Let $T:\mathbb{R}^n\to\mathbb{R}^m$ be any linear transformation. For elementary unit vectors,

e_1, e_2, ... , e_n of $\mathbb{R}^n$ and for any $x \in \mathbb{R}^n$, we have

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$$

and as $T(e_1)$, $T(e_2)$, ... , $T(e_n)$ are $m \times 1$ matrix, we can write them as

$$T(e_1) = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad T(e_2) = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \ldots \quad , \quad T(e_n) = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Therefore any linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ can be expressed as

$$\begin{aligned} T(x) &= x_1 T(e_1) + x_2 T(e_2) + \cdots + x_n T(e_n) \\ &= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}. \end{aligned} \tag{1}$$

Now let A be an $m \times n$ matrix which has $T(e_1)$, $T(e_2)$, ... , $T(e_n)$ as it's columns.

$$A = [\, T(e_1) : T(e_2) : \cdots : T(e_n) \,] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Then,

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

The above matrix $A = [a_{ij}]_{m \times n}$ is called the standard matrix of T and is denoted by $[T]$. Hence, the standard matrix of the linear transformation given by (1) can be found easily from the column vectors by substituting the elementary unit vectors to T in that order.

Theorem 6.1.2 [Standard Matrix]

If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then the standard matrix $A = [T]$ of T has the following relation for $\mathbf{x} \in \mathbb{R}^n$.

$$T(\mathbf{x}) = A\mathbf{x}, \quad \forall\, \mathbf{x} \in \mathbb{R}^n$$

where $A = [\, T(\mathbf{e}_1) : T(\mathbf{e}_2) : \cdots : T(\mathbf{e}_n) \,]$.

Example 6

For a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^4$, $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 2x_2 \\ -x_1 - x_2 \\ x_3 \\ x_1 + x_3 \end{bmatrix}$, by using the standard matrix of T, rewrite it as $T(\mathbf{x}) = A\mathbf{x}$.

Solution Let $A = \begin{bmatrix} 1 & 2 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$, which columns are $T(\mathbf{e}_i)$, then $T(\mathbf{x}) = A\mathbf{x}$ as

$$T(\mathbf{x}) = T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 2x_2 \\ -x_1 - x_2 \\ x_3 \\ x_1 + x_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = A\mathbf{x}. \quad \square$$

- http://matrix.skku.ac.kr/RPG_English/6-MA-standard-matrix.html

Sage http://sage.skku.edu

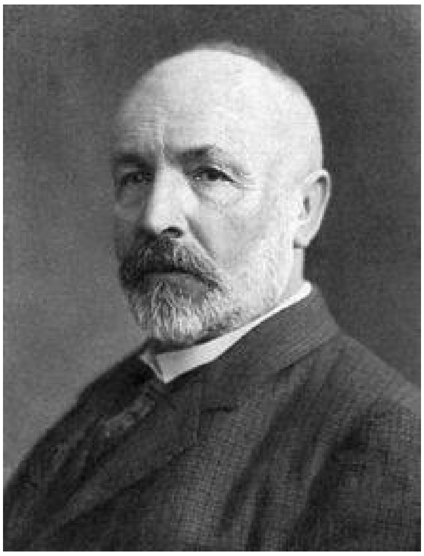
```
x, y, z = var('x y z')
h(x, y, z) = [x+2*y, -x-y, z, x+z]
T = linear_transformation(QQ^3, QQ^4, h)  # define linear transformation,
                                          # here scalar is rational numbers
C = T.matrix(side='right')                # standard matrix
x0 = vector(QQ, [2, -3, 3])
print C
print T.domain()                          # domain
print T.codomain()                        # codomain
print T(x0)                               # image
print C*x0                    # product of standard matrix and a vector
```

```
[ 1  2  0]
[-1 -1  0]
[ 0  0  1]
[ 1  0  1]
Vector space of dimension 3 over Rational Field
Vector space of dimension 4 over Rational Field
(-4, 1, 3, 5)
(-4, 1, 3, 5)
```
■

http://en.wikiquote.org/wiki/Georg_Cantor

Georg Ferdinand Ludwig Philipp Cantor (1845-1918)

"The essence of mathematics lies entirely in its freedom."
most famous as the creator of set theory, and of Cantor's theorem which implies the existence of an "infinity of infinities."

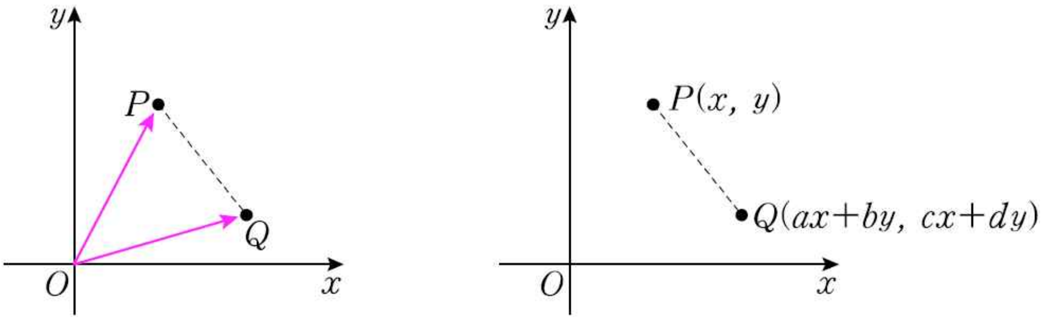


6.2 Geometric Meaning of Linear Transformations

- Reference video: http://http://youtu.be/cgySDj−OVIM, http://youtu.be/12WP−cb6Ymc
- Practice site: http://matrix.skku.ac.kr/knou−knowls/cla−week−8−Sec−6−2.html

In this section, we study **the geometric meaning of linear transformations**. For a given image, continuous showing of series of images with small variations makes a motion picture. Linear transformation can be applied to computer graphics and numerical algorithms, and it is an essential tool for many areas such as animation.

Linear Transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$



- A linear transformation $T: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} ax+by \\ cx+dy \end{bmatrix}$ moves a vector $\overrightarrow{OP} = (x, y)$ to an another vector $\overrightarrow{OQ} = (ax+by,\ cx+dy)$.

Example 1 [rotation, symmetry, orthogonal projection] We illustrate a few linear transformations on $\mathbb{R}^2$.

(1) $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation which rotates a vector in $\mathbb{R}^2$ counterclockwise by θ around the origin.

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

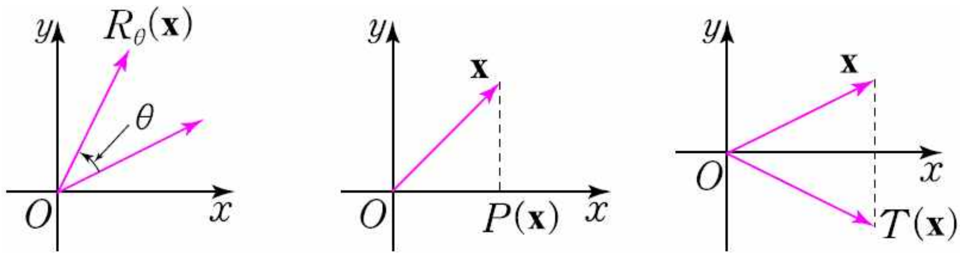
(2) An orthogonal projection $P: \mathbb{R}^2 \to \mathbb{R}^2$ on x-axis is a linear transformation.

$$P(\mathrm{x}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}$$

(3) A symmetric movement $T: \mathbb{R}^2 \to \mathbb{R}^2$ around x-axis is a linear transformation.

$$T(\mathbf{x}) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix}$$ ■

- http://matrix.skku.ac.kr/sglee/LT/11.swf



Example **2** Find the standard matrix A for a linear transformation which moves a point $P(x, y)$ in $\mathbb{R}^2$ to a symmetric image around the given line.

(1) y-axis (2) line $y = x$

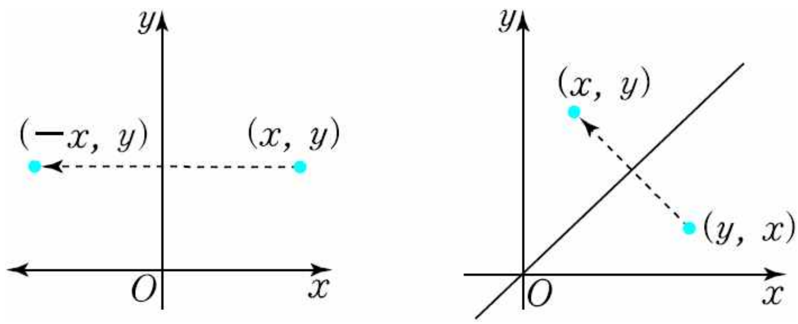
Solution

Symmetric (linear) transformation around y-axis and the line $y=x$ are given in the following figures.

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -x \\ y \end{bmatrix}, \qquad T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} y \\ x \end{bmatrix}$$

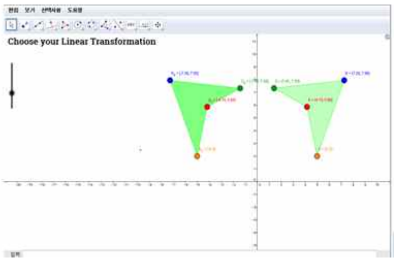
$$A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$ ■

- http://matrix.skku.ac.kr/sglee/LT/22.swf
- http://matrix.skku.ac.kr/sglee/LT/44.swf



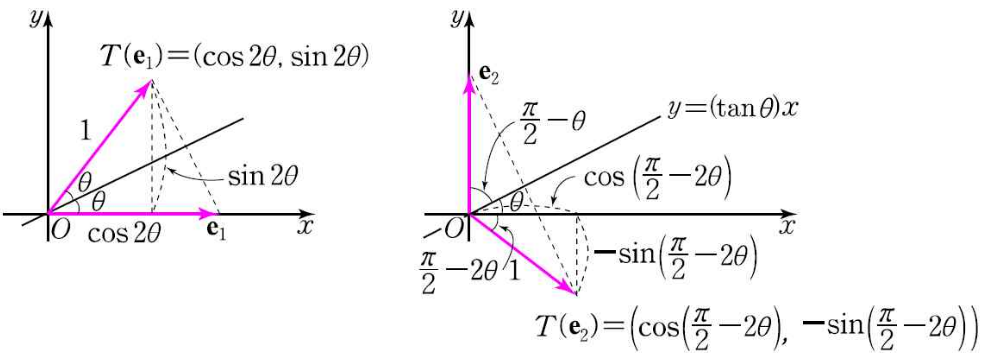
[Remark] **Simulation**

[linear transformation] http://www.geogebratube.org/student/m9703
[symmetric transformations and orthogonal projection transformations]
http://www.geogebratube.org/student/m9910
[rotation] http://www.geogebratube.org/student/m9702



Example 3 Linear transformation $T:\mathbb{R}^2 \to \mathbb{R}^2$ which moves any vector $\mathbf{x}=(x,y)$ in $\mathbb{R}^2$ to a symmetric image around a line, which passes through the origin with angle θ between the x-axis and the line, can be expressed by the following matrix presentation $H_\theta = [T(\mathbf{e}_1) : T(\mathbf{e}_2)]$.

$$H_\theta = [T(\mathbf{e}_1) : T(\mathbf{e}_2)] = \begin{bmatrix} \cos 2\theta & \cos\left(\frac{\pi}{2} - 2\theta\right) \\ \sin 2\theta & -\sin\left(\frac{\pi}{2} - 2\theta\right) \end{bmatrix} = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \quad ■$$



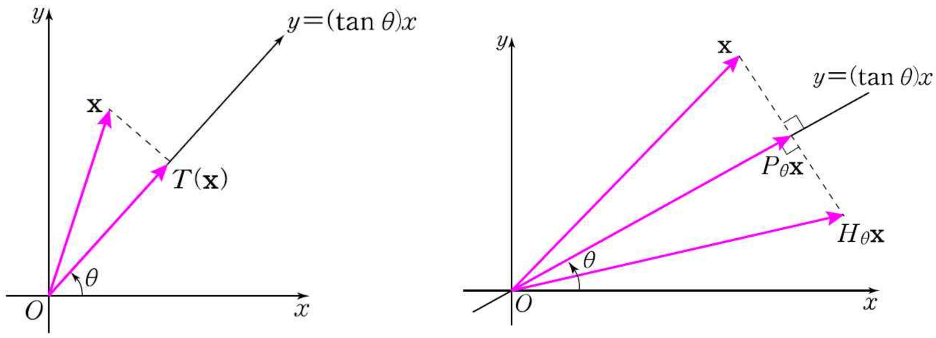
In **Example 3**, if $\theta=\frac{\pi}{4}$, $H_\theta = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, i.e. $T(x,y)=(y,x)$.

Example 4 As shown in the picture, let us define an orthogonal projection as a linear transformation (linear operator) $T:\mathbb{R}^2 \to \mathbb{R}^2$ which maps any vector $\mathbf{x}$ in $\mathbb{R}^2$ to the orthogonal projection on a line, which passes through the origin with angle θ between the x-axis and the line. Let us denote the standard matrix correspond to T is P_θ.

$$P_\theta \mathbf{x} - \mathbf{x} = \frac{1}{2}(H_\theta \mathbf{x} - \mathbf{x}) \quad \text{(the same direction with a half length)}$$

$$P_\theta \mathbf{x} = \frac{1}{2}H_\theta \mathbf{x} + \frac{1}{2}\mathbf{x} = \frac{1}{2}H_\theta \mathbf{x} + \frac{1}{2}I\mathbf{x} = \frac{1}{2}(H_\theta + I)\mathbf{x}$$

$$P_\theta = \frac{1}{2}(H_\theta + I) = \begin{bmatrix} \frac{1}{2}(1+\cos 2\theta) & \frac{1}{2}\sin 2\theta \\ \frac{1}{2}\sin 2\theta & \frac{1}{2}(1-\cos 2\theta) \end{bmatrix} = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix} \quad \blacksquare$$



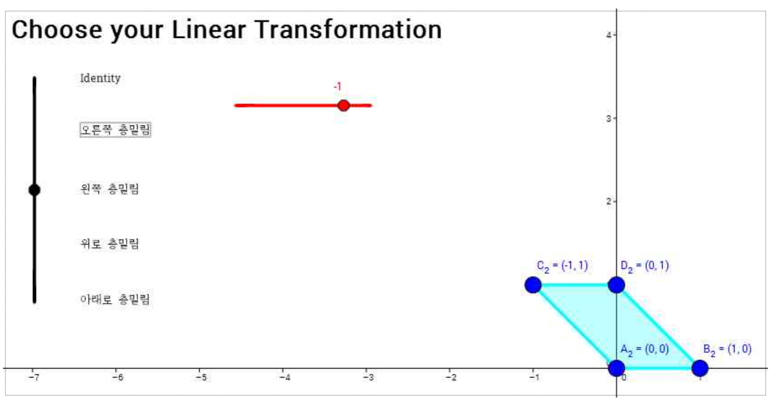
In Example 4, if $\theta = 0$, $P_\theta = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is a projection onto the $\mathbf{x}$-axis.

[Remark] shear transformations (computer simulation)

(1) $\begin{bmatrix} x \\ y \end{bmatrix} \to \begin{bmatrix} x+ky \\ y \end{bmatrix}$: shear transformation along the x-axis with scale k

(2) $\begin{bmatrix} x \\ y \end{bmatrix} \to \begin{bmatrix} x \\ kx+y \end{bmatrix}$: shear transformation along the y-axis with scale k

- http://www.geogebratube.org/student/m9912



Definition

A linear transformation $T: \mathbb{R}^n \to \mathbb{R}^n$, which preserve the magnitude (or length of a vector), $\|T(\mathbf{x})\| = \|\mathbf{x}\|$, is called Euclidean isometry.

Theorem 6.2.1

For a linear operator $T: \mathbb{R}^n \to \mathbb{R}^n$, the following statements are equivalent:

(1) $\|T(\mathbf{x})\| = \|\mathbf{x}\|$, $\mathbf{x} \in \mathbb{R}^n$ (isometry).

(2) $T(\mathbf{x}) \cdot T(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (preserve the inner product).

Definition

For a square matrix A, if $A^{-1} = A^T$ then A is called orthogonal matrix.

Example 5

For any real number θ, $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal matrix, and $Q^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$. ■

Example 6

Verify the following matrices are orthogonal matrix.

$$A = \begin{bmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix}, \quad B = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{bmatrix}$$

Solution

Verify $A^T A = I$, $B^T B = I$ by using the Sage. □

- http://matrix.skku.ac.kr/RPG_English/6-TF-orthogonal-matrix.html

Sage http://sage.skku.edu

```
A=matrix(QQ, 2, 2, [3/5, -4/5, 4/5, 3/5])
B=matrix(3, 3, [1/sqrt(3), 1/sqrt(2), 1/sqrt(6),
                1/sqrt(3), -1/sqrt(2), 1/sqrt(6),
                1/sqrt(3), 0, -2/sqrt(6)])
print A.transpose()*A                    # confirm the orthogonal matrix
print
print B.transpose()*B
```

```
[1 0]     [1 0 0]
[0 1]     [0 1 0]
          [0 0 1]
```

■

Theorem 6.2.2

For any $n \times n$ matrix A, the following statements are hold:

(1) The transpose of an orthogonal matrix is an orthogonal matrix.
(2) The inverse of an orthogonal matrix is an orthogonal matrix.
(3) The product of orthogonal matrices is an orthogonal matrix.
(4) If A is an orthogonal matrix, then $\det A = 1$ or -1.

Proof (1) and (2) are left as an exercise to the reader.

(3) If $A^{-1} = A^T$ and $B^{-1} = B^T$, then $(AB)^{-1} = B^{-1}A^{-1} = B^T A^T = (AB)^T$ and hence AB is an orthogonal matrix.

(4) Observe that $1 = \det I = \det(AA^T) = \det(A)\det(A^T) = (\det A)^2$

$\therefore\ \det A = 1$ or -1 . ■

Theorem 6.2.3

For any $n \times n$ matrix A, the following statements are equivalent:

(1) A is an orthogonal matrix.

(2) $\|A\mathrm{x}\| = \|\mathrm{x}\|$, $\mathrm{x} \in \mathbb{R}^n$.

(3) $A\mathrm{x} \cdot A\mathrm{y} = \mathrm{x} \cdot \mathrm{y}$, $\mathrm{x}, \mathrm{y} \in \mathbb{R}^n$.

(4) The row vectors of A are orthonormal.

(5) The column vectors of A are orthonormal.

Proof (1) $\Rightarrow$ (2): $\|A\mathrm{x}\|^2 = A\mathrm{x} \cdot A\mathrm{x} = \langle A\mathrm{x}, A\mathrm{x} \rangle = (A\mathrm{x})^T A\mathrm{x} = \mathrm{x}^T A^T A\mathrm{x}$

$$= \mathrm{x}^T A^{-1} A\mathrm{x} = \mathrm{x}^T \mathrm{x} = \langle \mathrm{x}, \mathrm{x} \rangle = \mathrm{x} \cdot \mathrm{x} = \|\mathrm{x}\|^2$$

(2) $\Rightarrow$ (3): $\| A(\mathrm{x}+\mathrm{y}) \|^2 = \| A\mathrm{x} + A\mathrm{y} \|^2 = \| A\mathrm{x} \|^2 + 2(A\mathrm{x} \cdot A\mathrm{y}) + \| A\mathrm{y} \|^2$

$$= \| \mathrm{x} \|^2 + 2(A\mathrm{x} \cdot A\mathrm{y}) + \| \mathrm{y} \|^2$$

and

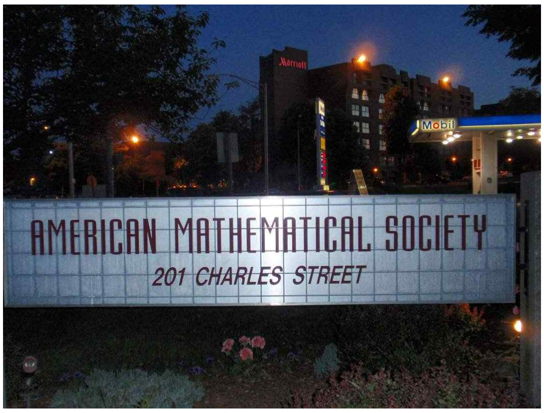
$$\| A(\mathrm{x}+\mathrm{y}) \|^2 = \| \mathrm{x}+\mathrm{y} \|^2 = \| \mathrm{x} \|^2 + 2(\mathrm{x} \cdot \mathrm{y}) + \| \mathrm{y} \|^2.$$

Hence $A\mathrm{x} \cdot A\mathrm{y} = \mathrm{x} \cdot \mathrm{y}$.

(3) $\Rightarrow$ (1): $\forall i$, $\mathrm{e}_j{}^T A^T A \mathrm{e}_i = \mathrm{e}_i \cdot \mathrm{e}_j = \langle \mathrm{e}_i, \mathrm{e}_j \rangle = \mathrm{e}_j{}^T \mathrm{e}_i = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

$$\Rightarrow (A^T A)_{ij} = \sigma_{ij} \qquad \therefore \; A^T A = I_n$$

We skip the detailed proof of (4) and (5) as we can get the result easily from the definition of the orthogonal matrix, $A^T A = I = AA^T$, and (1). ■



[The headquarter of American Mathematical Society, Providence, RI, USA]
http://www.ams.org

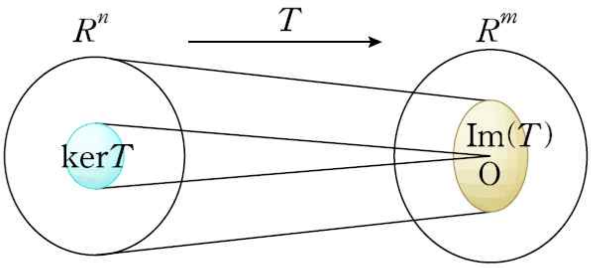
6.3 Kernel and Range

- Reference video: http://youtu.be/9YciT9Bb2B0, http://youtu.be/H-P4lDgruCc
- Practice site: http://matrix.skku.ac.kr/knou-knowls/cla-week-8-Sec-6-3.html

We will show that the subset of a domain $\mathbb{R}^n$, which maps to zero vector by a linear transformation, becomes a subspace. We will also show the set of images under any linear transformation forms a subspace in the co-domain. Finally, we introduce the concept of isomorphism.

Definition

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. The set of all vectors in $\mathbb{R}^n$, whose image becomes $\mathbf{0}$ by T, is called kernel of T and is denoted by $\ker T$. That is, $\ker T = \{\mathbf{v} \in \mathbb{R}^n \mid T(\mathbf{v}) = \mathbf{0}\}$.



Example 1 Find the $\ker T$ for a linear transformation $T: \mathbb{R}^2 \to \mathbb{R}^2$, where $T(x, y) = (x - y, 0)$.

Solution

$\ker T = \{(x, y) \in \mathbb{R}^2 \mid (x - y, 0) = (0, 0)\} = \{(x, y) \in \mathbb{R}^2 \mid y = x\}$. ■

Example 2 Find the $\ker T$ for a linear transformation $T: \mathbb{R}^4 \to \mathbb{R}^4$, where $T(x_1, x_2, x_3, x_4) = (0, x_1, x_2, x_3)$.

Solution

For any $\mathbf{x} = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4$,

$$T(x_1,x_2,x_3,x_4) = (0,x_1,x_2,x_3) = \mathbf{0} \quad \Leftrightarrow \quad x_i = 0,\ i = 1,2,3$$

and hence, $\ker T = \{(0,0,0,x_4) \mid x_4 \in \mathbb{R}\}$. ■

Definition

For a transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, if $T(\mathrm{u}) = T(\mathrm{v}) \Rightarrow \mathrm{u} = \mathrm{v}$, then it is called **one-to-one** (injective).

Definition

For a transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, if there exist $\mathrm{v} \in \mathbb{R}^n$ for any given $\mathrm{w} \in \mathbb{R}^m$, such that $T(\mathrm{v}) = \mathrm{w}$, then it is called **onto** (surjective).

Theorem 6.3.1

Let $\mathbb{R}^n$ and $\mathbb{R}^m$ are vector spaces and $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Then T is one-to-one if and only if $\ker T = \mathbf{0}$.

Proof $(\Rightarrow)$ As $\forall\, \mathrm{v} \in \ker T$, $T(\mathrm{v}) = \mathbf{0} = T(\mathbf{0})$ and T is one-to-one,

$\Rightarrow \quad \mathrm{v} = \mathbf{0} \quad \therefore \quad \ker T = \{\mathbf{0}\}$

$(\Leftarrow)$ $T(\mathrm{v}_1) = T(\mathrm{v}_2) \Rightarrow \mathbf{0} = T(\mathrm{v}_1) - T(\mathrm{v}_2) = T(\mathrm{v}_1 - \mathrm{v}_2)$

$\Rightarrow \mathrm{v}_1 - \mathrm{v}_2 \in \ker T = \{\mathbf{0}\} \Rightarrow \mathrm{v}_1 = \mathrm{v}_2$

$\therefore$ T is one-to-one. ■

Example 3 Let us define a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ as $T(x, y) = (y, x)$. Is Tan one-to-one?

Solution

As $\ker T = \{\mathrm{x} \in \mathbb{R}^2 \mid T(\mathrm{x}) = \mathbf{0}\} = \{(x, y) \mid T(x, y) = (y, x) = (0, 0)\}$, the only element in this set is $(x, y) = (0, 0)$. Hence $\ker T = \{(0, 0)\}$, and T is one-to-one. ■

Example 4 Is a linear transformation $T: \mathbb{R}^3 \to \mathbb{R}^3$ one-to-one if it is defined as $T(x, y, z) = (x + 2y - z, y + z, x + y - 2z)$?

Solution

Since $T(x, y, z) = (0, 0, 0)$ $\Leftrightarrow$ $\begin{cases} x + 2y - z = 0 \\ \quad\; y + z = 0 \\ x + y - 2z = 0 \end{cases}$ $\Leftrightarrow$ $\begin{cases} x + 2y - z = 0 \\ \quad\; y + z = 0 \end{cases}$,

the system of linear equations has infinitely many solutions. Hence, $\ker T \neq \{0\}$ and by **theorem 6.3.1**, T is not one-to-one. □

Sage http://sage.skku.edu

① verify linear transformations' one-to-one

```
U = QQ^3                                # vector space
x, y, z = var('x, y, z')
h(x, y, z) = [x+2*y-z, y+z, x+y-2*z]
T = linear_transformation(U, U, h)      # generate a linear transformation
print T.is_injective()                  # check the one-to-one
```

```
False
```

② Find Kernel of linear transformation

```
T.kernel()                              # verify by finding kernel
```

```
Vector space of degree 3 and dimension 1 over Rational Field
Basis matrix:
[   1 -1/3  1/3]                  # kernel = span( [1, -1/3, 1/3] ).
```
■

Let A be an $m \times n$ matrix. If we define a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ as $T(\mathbf{x}) = A\mathbf{x}$, then $\ker T$ is a solution space of the system of linear equations $A\mathbf{x} = \mathbf{0}$.

Theorem 6.3.2

Let $\mathbb{R}^n$, $\mathbb{R}^m$ are vector spaces and $T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Then $\ker T$ is a subspace of $\mathbb{R}^n$. Hence $\ker T$ is called kernel (subspace).

Example 5

Find the kernel of a $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

Sage http://sage.skku.edu

```
A = matrix(2, 2, [1, 1, 1, -1])
print A.right_kernel()                    # kernel of A
```

```
Free module of degree 2 and rank 0 over Integer Ring
Echelon basis matrix:
[]                                # kernel has only 0.
```

■

Definition [Isomorphism]

For a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$, the set of all $T(\mathbf{v})$ for $\mathbf{v} \in \mathbb{R}^n$, is called **range** of T and is denoted by $\operatorname{Im} T$. That is,

$$\operatorname{Im} T = \{T(\mathbf{v}) \in \mathbb{R}^m \mid \mathbf{v} \in \mathbb{R}^n\} \subset \mathbb{R}^m.$$

Especially, if $\operatorname{Im} T = \mathbb{R}^m$ then T is called surjective or onto. If a linear transformation T is one-to-one and onto, then $n = m$ and T is called an **isomorphism** from $\mathbb{R}^n$ to $\mathbb{R}^n$.

Example 6

Find the range of the linear transformation $T(x, y) = (x - y, 0)$.

Solution

$\operatorname{Im} T = \{T(x, y) \mid (x, y) \in \mathbb{R}^2\} = \{(x - y, 0) \mid (x, y) \in \mathbb{R}^2\} = \{(a, 0) \mid a \in R\}$.

Note that, $\operatorname{Im} T \neq \mathbb{R}^2 \Rightarrow T$ is not surjective. ■

- T is not isomorphism as it is not surjective.

Example **7** Let $W_1 = \{(x_1, x_2, 0, 0) \mid x_1, x_2 \in \mathbb{R}\}$ and $W_2 = \{(0, 0, x_3, x_4) \mid x_3, x_4 \in \mathbb{R}\}$. It is easy to see that both W_1 and W_2 are subspaces of R^4. If we define $T: W_1 \to W_2$ as following linear transformation,

$$T(x, y, 0, 0) = (0, 0, x, y)$$

then T is both one-to-one and onto, and hence it is isomorphism. ■

Theorem 6.3.3

For a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$, $\operatorname{Im} T$ is a subspace of $\mathbb{R}^m$.

Proof $\forall\, \mathrm{w}_1, \mathrm{w}_2 \in \operatorname{Im} T,\ \exists\, \mathrm{v}_1, \mathrm{v}_2 \in \mathbb{R}^n \ni T(\mathrm{v}_1) = \mathrm{w}_1,\ T(\mathrm{v}_2) = \mathrm{w}_2$

$\Rightarrow\ \mathrm{w}_1 + \mathrm{w}_2 = T(\mathrm{v}_1) + T(\mathrm{v}_2) = T(\mathrm{v}_1 + \mathrm{v}_2)$

$\Rightarrow\ \exists\, \mathrm{v}_1 + \mathrm{v}_2 \in \mathbb{R}^n \ni T(\mathrm{v}_1 + \mathrm{v}_2) = \mathrm{w}_1 + \mathrm{w}_2 \in \mathbb{R}^m \quad \therefore\ \mathrm{w}_1 + \mathrm{w}_2 \in \operatorname{Im} T$

$\forall\, k \in R,\ k\mathrm{w}_1 = kT(\mathrm{v}_1) = T(k\mathrm{v}_1)$

$\Rightarrow\ \exists\, k\mathrm{v}_1 \in R^n \ni T(k\mathrm{v}_1) = k\mathrm{w}_1 \in \mathbb{R}^m \quad \therefore\ k\mathrm{w}_1 \in \operatorname{Im} T$

$\therefore$ $\operatorname{Im}(T)$ is a subspace of $\mathbb{R}^m$. ■

Example **8** Let A be an $m \times n$ matrix, if we define a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ as $T(\mathrm{x}) = A\mathrm{x}$, then $\operatorname{Im} T$ is a column space of A.

Solution

Let $A = [A^{(1)} : A^{(2)} : \cdots : A^{(n)}]$, that is, $A^{(i)}$ be an $m \times n$ matrix A's ith column vector. Then for any vector $\mathrm{x} = (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n$,

$$A\mathrm{x} = [A^{(1)} A^{(2)} \cdots A^{(n)}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 A^{(1)} + x_2 A^{(2)} + \cdots + x_n A^{(n)}.$$

That is, any image can be expressed as a linear combination of column vectors of A.

$\therefore$ $\operatorname{Im} T = \{A\mathrm{x} \mid \mathrm{x} \in \mathbb{R}^n\} = \langle A^{(1)}, A^{(2)}, \dots, A^{(n)} \rangle$ ■

Theorem 6.3.4

For a linear transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by a matrix $A = [a_{ij}]_{m \times n}$ satisfies the following two properties.

(1) T_A is one-to-one. ⇔ column vectors of A are linearly independent.

(2) T_A is onto. ⇔ row vectors of A are linearly independent.

Proof (1) T_A in one-to-one ⇔ $\ker T_A = \{\mathbf{0}\}$

⇔ There is a unique $\mathbf{x} = \mathbf{0} \in \mathbb{R}^n$ which satisfies $A\mathbf{x} = \mathbf{0}$.

⇔ n column vectors of A are linearly independent.

(2) T_A is onto ⇔ $\operatorname{Im} T_A = \mathbb{R}^m$

⇔ For A's column vectors $A^{(i)}$,

$\mathbb{R}^m = \operatorname{Im} T = \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\} = < A^{(1)}, A^{(2)}, \ldots, A^{(n)} >$

⇔ In RREF(A), the number of leading ones is m.

⇔ row rank of A is m.

⇔ m row vectors of A are linearly independent. ■

Example 9 Verify the following by using the Sage.

(1) Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$. $T_A : \mathbb{R}^2 \to \mathbb{R}^3$ is one-to-one but not onto.

Sage http://sage.skku.edu

① define a linear transformation

```
U = QQ^2
V = QQ^3
A = matrix(QQ, [[1, 0], [0, 1], [0, 0]])
T = linear_transformation(U, V, A, side='right')    # linear transformation
print T
```

```
Vector space morphism represented by the matrix:
[1 0 0]
[0 1 0]
Domain: Vector space of dimension 2 over Rational Field
Codomain: Vector space of dimension 3 over Rational Field
```

② check the surjectivity (onto)

```
print T.image()              # generate the range
print T.is_surjective()      # check the surjectivity (onto)
```

```
Vector space of degree 3 and dimension 2 over Rational Field
Basis matrix:
[1 0 0]
[0 1 0]
False
```

③ check the injectivity (one-to-one)

```
print T.kernel()             # generate the kernel
print T.is_injective()       # check the injectivity (one-to-one)
```

```
Vector space of degree 2 and dimension 0 over Rational Field
Basis matrix:
[]
True
```

(2) Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. $T_A : \mathbb{R}^3 \to \mathbb{R}^2$ is onto but not one-to-one.

Sage http://sage.skku.edu

① define a linear transformation

```
U = QQ^3
V = QQ^2
A = matrix(QQ, [[1, 0, 0], [0, 1, 0]])
T = linear_transformation(U, V, A, side='right')     # linear transformation
print T
```

```
Vector space morphism represented by the matrix:
[1 0]
[0 1]
[0 0]
Domain: Vector space of dimension 3 over Rational Field
odomain: Vector space of dimension 2 over Rational Field
```

② check the surjectivity (onto)

```
print T.image()                # generate the range
print T.is_surjective()         # check the surjectivity (onto)
```

```
Vector space of degree 2 and dimension 2 over Rational Field
Basis matrix:
[1 0]
[0 1]
True
```

③ check the injectivity (one-to-one)

```
print T.kernel()                 # generate the kernel
print T.is_injective()           # check the injectivity (one-to-one)
```

```
Vector space of degree 3 and dimension 1 over Rational Field
Basis matrix:
[0 0 1]
False
```

■

Theorem 6.3.5

Let $A = [a_{ij}]_{n \times n}$ be an $n \times n$ matrix. If $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation, T_A is one-to-one if and only if T_A is onto.

Proof T_A is one-to-one $\Leftrightarrow$ $\ker T_A = \{\mathbf{0}\}$

$\Leftrightarrow$ There is a unique $\mathbf{x} = \mathbf{0} \in \mathbb{R}^n$ which satisfies $A\mathbf{x} = \mathbf{0}$.

$\Leftrightarrow$ In A's RREF, number of leading ones is n.

$\Leftrightarrow$ For A's column vectors $A^{(i)}$,

$\operatorname{Im} T = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\} = < A^{(1)}, A^{(2)}, \ldots, A^{(n)} > = \mathbb{R}^n$

$\Leftrightarrow$ $\operatorname{Im} T_A = \mathbb{R}^n$ $\Leftrightarrow$ T_A is onto. ■

Equivalence Theorem of Invertible Matrix

Theorem 6.3.6 [Equivalence Theorem of Invertible Matrix]

Let A be an $n \times n$ matrix, the following statements are all equivalent.

(1) column vectors of A are linearly independent.
(2) row vectors of A are linearly independent.
(3) $A\mathbf{x} = \mathbf{0}$ has only trivial solution $\mathbf{x} = \mathbf{0}$.
(4) For any $n \times 1$ vector $\mathbf{b}$, $A\mathbf{x} = \mathbf{b}$ has a unique solution.
(5) A and I_n are column equivalent.
(6) A is invertible.
(7) $\det(A) \neq 0$
(8) $\lambda = 0$ is not an eigenvalue of A.
(9) T_A is one-to-one.
(10) T_A is onto.

Country	Team size			P1	P2	P3	P4	P5	P6	Total	Rank
	All	M	F								
Republic of Korea	6	6		42	42	21	39	42	23	209	1
People's Republic of China	6	6		42	40	14	31	38	30	195	2
United States of America	6	6		42	40	33	38	23	18	194	3
Russian Federation	6	6		42	35	21	41	29	9	177	4
Canada	6	6		42	32	9	39	24	13	159	5
Thailand	6	6		42	42	4	39	30	2	159	5
Singapore	6	6		42	35	11	32	27	7	154	7
Islamic Republic of Iran	6	5	1	42	29	6	39	34	1	151	8
Vietnam	6	6		42	36	4	31	35	0	148	9
Romania	6	5	1	40	36	7	36	20	5	144	10

[Ranking of International Math Olympiad 2012]

https://www.imo-official.org/results.aspx

6.4 Composition of Linear Transformations and Invertibility

- Reference video: http://youtu.be/EOlq4LouGao http://youtu.be/qfAmNsdlPxc
- Practice site: http://matrix.skku.ac.kr/knou-knowls/cla-week-8-Sec-6-4.html

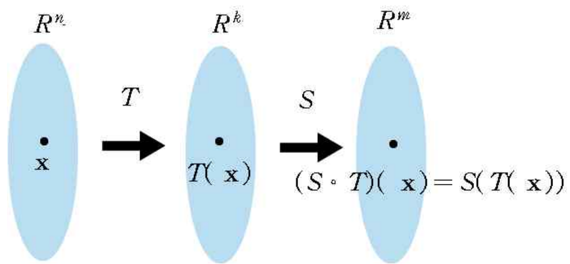
In this section, we study the composition of two or more linear transformations as continuous product of matrices. We also study the geometric properties of linear transformation by connecting inverse functions and inverse matrices.

Theorem 6.4.1 [Composition of Functions]

If both $T: \mathbb{R}^n \to \mathbb{R}^k$ and $S: \mathbb{R}^k \to \mathbb{R}^m$ are linear transformations, then the composition

$$S \circ T: \mathbb{R}^n \to \mathbb{R}^m$$

is also a linear transformation.



Theorem 6.4.2

For linear transformations $T: \mathbb{R}^n \to \mathbb{R}^k$ and $S: \mathbb{R}^k \to \mathbb{R}^m$,

(1) $S \circ T$ is one-to-one implies T is one-to-one.

(2) $S \circ T$ is onto implies S is onto.

Proof (1) If $T(\mathrm{v}_1) = T(\mathrm{v}_2)$, for $\mathrm{v}_1, \mathrm{v}_2 \in \mathbb{R}^n$, then $S(T(\mathrm{v}_1)) = S(T(\mathrm{v}_2))$.

$\Rightarrow$ $(S \circ T)(\mathrm{v}_1) = (S \circ T)(\mathrm{v}_2)$ $\Rightarrow$ $\mathrm{v}_1 = \mathrm{v}_2$ ($\because$ $S \circ T$ is one-to-one)

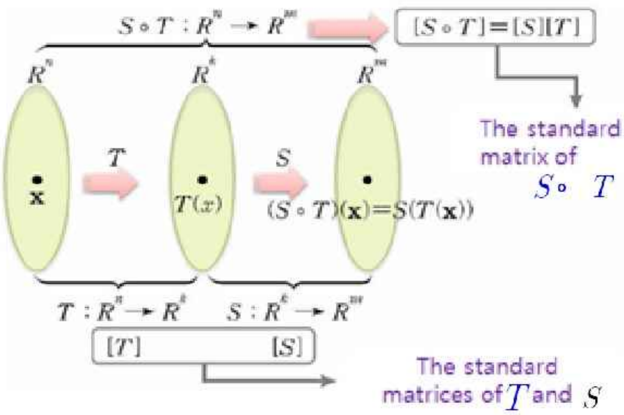
$\therefore$ T is one-to-one.

(2) If $S \circ T$ is onto, then for $\forall\, \mathrm{z} \in \mathbb{R}^m$, there exist $\mathrm{v} \in \mathbb{R}^n$ such that $(S \circ T)(\mathrm{v}) = \mathrm{z}$. That is, there exist $\mathrm{v} \in \mathbb{R}^n$ which satisfy $S(T(\mathrm{v})) = \mathrm{z}$. Since $T(\mathrm{v}) = \mathrm{w} \in \mathbb{R}^k$, there exist $\mathrm{w} \in \mathbb{R}^k$ such that $S(\mathrm{w}) = \mathrm{z}$.

$\therefore$ S is onto ■

For the case of composition of two linear transformations, the corresponding standard matrix is the product of two standard matrices from each linear transformation.

- That is, let $T:\mathbb{R}^n \to \mathbb{R}^k$, $S:\mathbb{R}^k \to \mathbb{R}^m$ and $T:\mathbb{R}^n \to \mathbb{R}^k$ has a standard matrix $[T]$, $S:\mathbb{R}^k \to \mathbb{R}^m$ has standard matrix $[S]$. Then the linear transformation $S\circ T:\mathbb{R}^n \to \mathbb{R}^m$ has the standard matrix $[S\circ T]=[S][T]$.



Let the standard matrix of a linear transformation T be A. If an inverse transformation T^{-1} exist, then the standard matrix of T^{-1} is the inverse of the matrix A.

Example 1

Let $T, S:\mathbb{R}^2 \to \mathbb{R}^2$ are linear transformations which rotate θ_1 and θ_2 (counterclockwise) respectively around the origin. The corresponding standard matrices are as follows.

$$[T]=\begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix},\ [S]=\begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix}$$

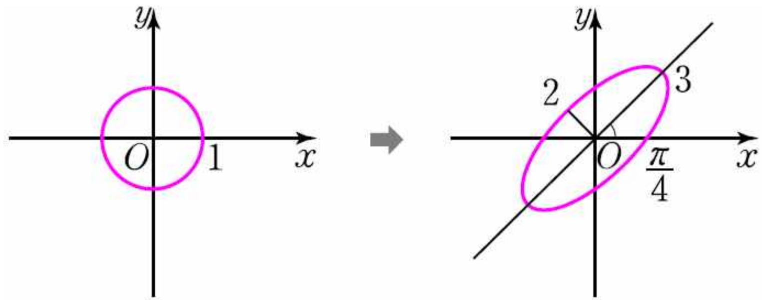
As the composition of these two transformations rotates $\theta_1+\theta_2$ around the origin, $R=S\circ T$'s standard matrix is as follows.

$$[R]=\begin{bmatrix} \cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2) \\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) \end{bmatrix}.$$

Also the product of standard matrices of T and S are as follows.

$$[S][T]=\begin{bmatrix}\cos\theta_2 & -\sin\theta_2\\ \sin\theta_2 & \cos\theta_2\end{bmatrix}\begin{bmatrix}\cos\theta_1 & -\sin\theta_1\\ \sin\theta_1 & \cos\theta_1\end{bmatrix}$$
$$=\begin{bmatrix}\cos\theta_2\cos\theta_1-\sin\theta_2\sin\theta_1 & -\cos\theta_2\sin\theta_1-\sin\theta_2\cos\theta_1\\ \sin\theta_2\cos\theta_1+\cos\theta_2\sin\theta_1 & -\sin\theta_2\sin\theta_1+\cos\theta_2\cos\theta_1\end{bmatrix}$$
$$=\begin{bmatrix}\cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2)\\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2)\end{bmatrix}=[R]=[S\circ\ T].$$ ■

Example 2 As shown in the picture, find a matrix transformation which transform a circle with radius 1 to the given ellipse.



Solution

First we find a transformation which expands 3 times around the x-axis, and expands 2 times around the y-axis. Then take a transformation which rotates $\frac{\pi}{4}$ around the origin. The first transformation T_1 is $T_1(x, y)=(3x, 2y)$, and hence the standard matrices for T_1 and the rotation transformation T_2 are

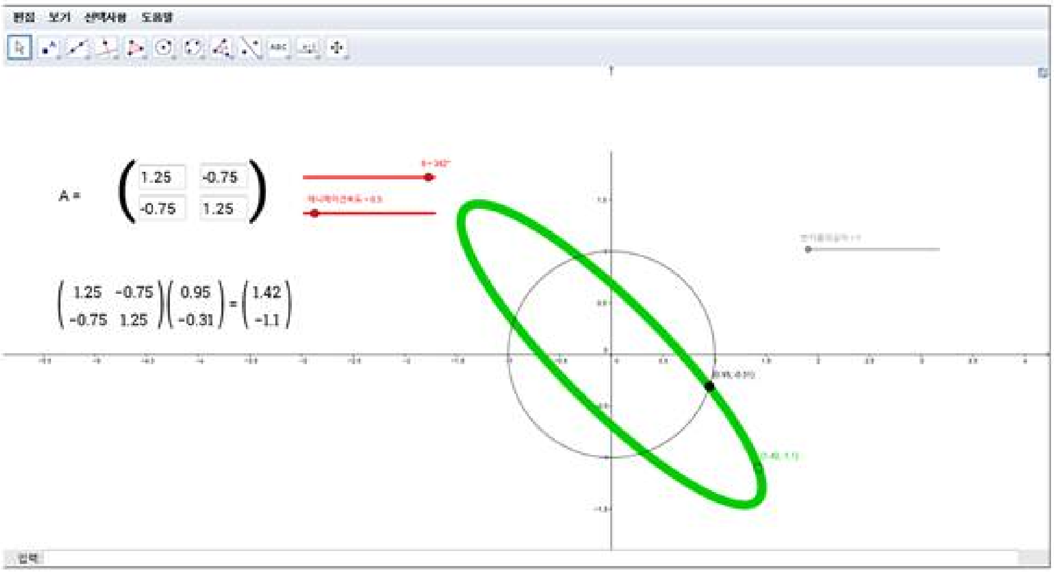
$$[T_1]=\begin{bmatrix}3&0\\0&2\end{bmatrix},\quad [T_2]=\begin{bmatrix}\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{bmatrix}.$$

Therefore, the standard matrix for the composition is the product of two standard matrices.

$$[T_2][T_1]=\begin{bmatrix}\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{bmatrix}\begin{bmatrix}3&0\\0&2\end{bmatrix}=\begin{bmatrix}\frac{3\sqrt{2}}{2} & -\sqrt{2}\\ \frac{3\sqrt{2}}{2} & \sqrt{2}\end{bmatrix}.$$ ■

[Remark] **Computer simulation**

[Matrix Transformation] http://www.geogebratube.org/student/m57556



- Similarly a composition of three or more linear transformations, the standard matrix of the composition is the product of each standard matrix in that operation order.

Theorem 6.4.3

A function $f : X \to Y$ is invertible if and only if f is one-to-one and onto.

Theorem 6.4.4

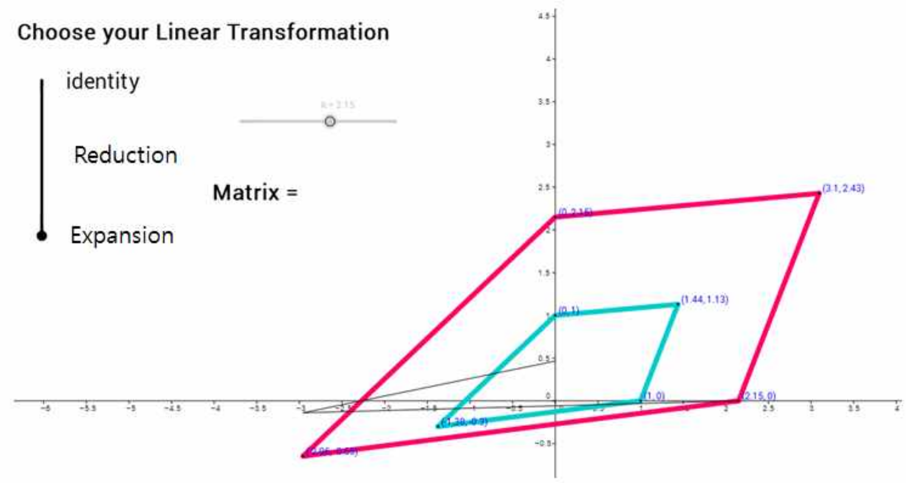
If a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is invertible, then $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is also a linear transformation.

- Inverse transformation of composition of transformation: $(S \circ T)^{-1} = T^{-1} \circ S^{-1}$

$$[S \circ T]^{-1} = ([S][T])^{-1} = [T]^{-1}[S]^{-1}$$

[Remark] Computer simulation

[shrink transformation and expand transformation]
http://www.geogebratube.org/student/m11366



“All human knowledge begins with intuitions, proceeds from thence to concepts, and ends with ideas.”

Immanuel Kant (1724-1804) is one of the most influential philosophers in the history of Western philosophy. His contributions to metaphysics, epistemology, ethics, and aesthetics have had a profound impact on almost every philosophical movement that followed him.

6.5 *Computer Graphics with Sage

- Reference video: http://youtu.be/VV5zzeYipZs
- Practice site: http://matrix.skku.ac.kr/Lab-Book/Sage-Lab-Manual-2.htm
  http://matrix.skku.ac.kr/Big-LA/LA-Big-Book-CG.htm

Computer graphics plays a key role in automotive design, flight simulation, and game industry. For example, a 3 dimensional object, such as automobile, its data (coordinates of points) can be described as a matrix. If we transform the location of these points, we can redraw the transformed object from the points which are newly generated. If this transformation is linear, we can easily obtain the transformed data by matrix multiplication. In this section, we review several geometric transformations which are used in computer graphics.

Geometric meaning of Linear Transformation 1 (Linear Transformation of Polygon's Image)

By using the Sage, draw a triangle with three vertices $(0,0)$, $(0,3)$, and $(3,0)$, a triangle expanded twice, a figure by a shear transformation along the x-axis with scale 1, and a triangle which is rotated counterclockwise by $\frac{\pi}{3}$.

- First of all, in order to define the above linear transformations, we input the following linear transformations by using matrix.

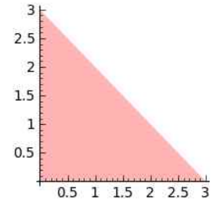
```
def matrix_transformation(A, L):
    n=matrix(L).nrows()     # list L's number of elements
    L2=[[0,0] for i in range(n)]    # define a new list L2
    for i in range(n):
        L2[i]=list(A*vector(L[i]))    # L2=A*L
    return L2    # return L2
print "The matrix_transformation function is activated"#confirm whether it is applied
```

- Then, we define appropriate standard matrices to fit the problems' condition.

```
A=matrix([[2,0], [0,2]])   # Expanding twice of given image
B=matrix([[1,1], [0,1]])   # shear transformation along the x-axis with scale 1
C=matrix([[cos(pi/3), -sin(pi/3)], [sin(pi/3), cos(pi/3)]])
                                       # rotate counterclockwise the given image by π/3
```

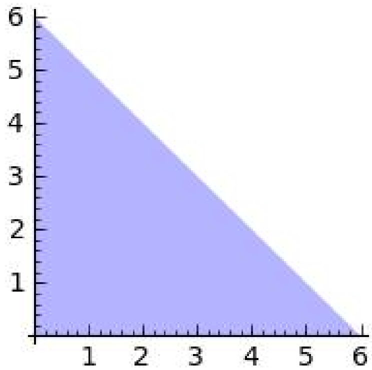
● Draw a triangle which has three vertices $(0, 0)$, $(0, 3)$, $(3, 0)$ by using ploygon.

```
L1=list( [ [0,0], [0,3], [3,0] ])   # input three vertices
SL1=polygon(L1, alpha=0.3, rgbcolor=(1,0,0))   # draw a polygon which passes
                                                 through the given three points
SL1.show(aspect_ratio=1, figsize=3)
```



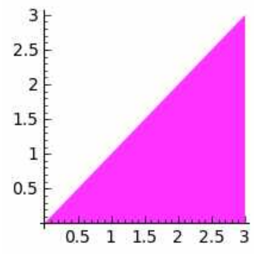
● Draw a twice expanded triangle from the given triangle.

```
L2=matrix_transformation(A, L1)   # find new three points by a linear transformation
SL2=polygon(L2, alpha=0.8, rgbcolor=(0,0,1))   # draw a polygon which passes
                                                through the given three points
SL2.show(aspect_ratio=1, figsize=3)
```



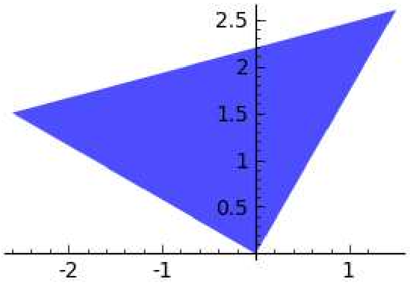
● Draw a shear transformed figure along the x-axis with scale 1 from the given triangle.

```
L3=matrix_transformation(B, L1)   # find new three points by a linear transformation
SL3=polygon(L3, alpha=0.8, rgbcolor=(1,0,1))   # draw a figure which passes
                                                through the given three points
SL3.show(aspect_ratio=1, figsize=3)
```



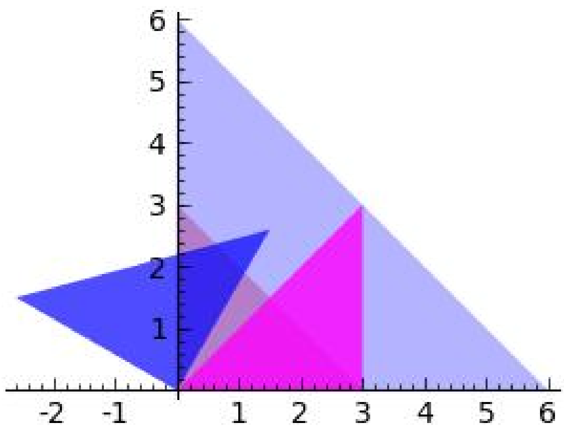
- Draw a figure which is rotated counterclockwise by $\frac{\pi}{3}$ from the given triangle.

```
L4=matrix_transformation(C, L1) # find new three points by a linear transformation
SL4=polygon(L4, alpha=0.4, rgbcolor=(0,0,1))    # draw a figure which passes
                                                  through the given three points
SL4.show(aspect_ratio=1, figsize=3)
```



- Show the above four figures in the same frame.

```
(SL1+SL2+SL3+SL4).show(aspect_ratio=1, figsize=3)
```



Geometric meaning of Linear Transformation 2 (Linear Transformation of Line's Image)

Draw the alphabet letter S on the plane. Then draw figures which expands the original figure twice, sheer transforms along the x-axis with scale 1, and rotates counterclockwise by $\frac{\pi}{3}$.

- First of all, in order to define the above linear transformations, we input the following linear transformations by using matrix.

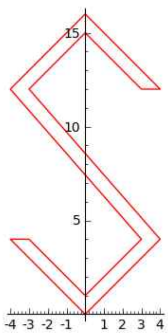
```
def matrix_transformation(A, L):
    n=matrix(L).nrows()    # list L's number of elements
    L2=[[0,0] for i in range(n)]    # define a new list L2
    for i in range(n):
        L2[i]=list(A*vector(L[i]))    # L2=A*L
    return L2    # return L2
print "The matrix_transformation function is activated"#confirm whether it is applied
```

- Then, we define appropriate standard matrices to fit the problems' condition.

```
A=matrix([[2,0], [0,2]])    # Expanding twice of given image
B=matrix([[1,1], [0,1]])    # shear transformation along the x-axis with scale 1
C=matrix([[cos(pi/3), -sin(pi/3)], [sin(pi/3), cos(pi/3)]])
                            # rotate counterclockwise the given image by π/3
```

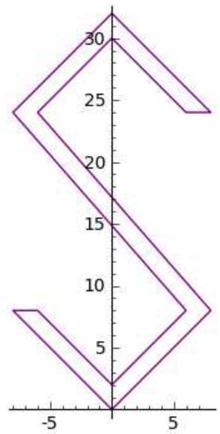
- Draw an alphabet letter S by using the line function.

```
L1=list( [ [0,0], [4,4], [-3,12], [0,15], [3,12], [4,12], [0,16], [-4,12], [3,4],
           [0,1], [-3,4], [-4,4], [0,0] ]) # input the data which compose letter S
SL1=line(L1, color="red")   # draw a figure which passes through the given points
SL1.show(aspect_ratio=1, figsize=5)
```



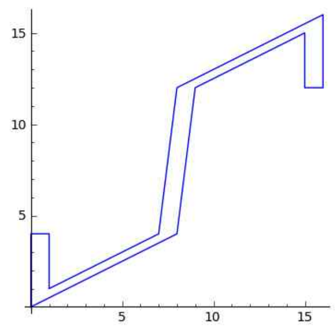
● Draw a twice expanded letter S from the given figure.

```
L2=matrix_transformation(A, L1)      # compute new points' coordinates by a linear transformation
SL2=line(L2, color="purple") #draw a figure which passes through the given points
SL2.show(aspect_ratio=1, figsize=5)
```



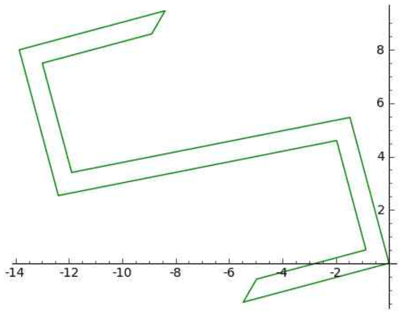
● Draw a sheer transformed figure along the x-axis with scale 1 from the given S.

```
L3=matrix_transformation(B, L1)    # compute new points' coordinates by a linear transformation
SL3=line(L3, color="blue")  # draw a figure which passes through the given points
SL3.show(aspect_ratio=1, figsize=5)
```



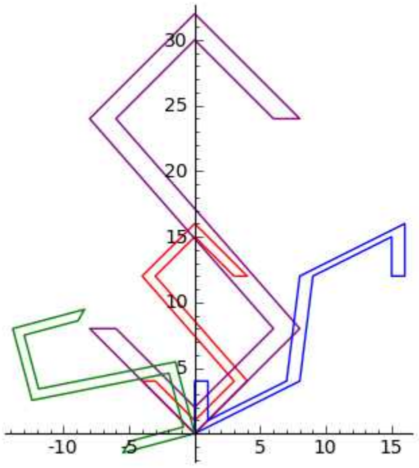
● Draw a figure which is rotated counterclockwise by $\frac{\pi}{3}$ from the given letter S.

```
L4=matrix_transformation(C, L1)    # compute new points' coordinates by a linear transformation
SL4=line(L4, color="green")   #draw a figure which passes through the given points
SL4.show(aspect_ratio=1, figsize=5)
```



- Show the above four figures in the same frame.

```
(SL1+SL2+SL3+SL4).show(aspect_ratio=1, figsize=5)
```



http://modular.math.washington.edu/
[William Stein : The first Sage developer]

[Sage developer group]

[Sage code developers: Linear Algebra]

Chapter 6 Exercises

- http://matrix.skku.ac.kr/LA-Lab/index.htm
- http://matrix.skku.ac.kr/knou-knowls/cla-sage-reference.htm

Problem 1 Verify that $T: \mathbb{R}^3 \to \mathbb{R}^3$, where $T(x_1, x_2, x_3) = (x_1, x_1 + x_2, x_1 + x_2 + x_3)$, is a linear transformation and find $T(\mathbf{x})$ for $\mathbf{x} = (1, 2, 3)$.

Solution The map T can be written as a matrix transformation, $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. So it is a linear transformation. $T(\mathbf{x}) = T\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix}$. ■

Problem 2 Find the standard matrix $[T]$ for $T(x, y, z) = (x - y, y - z, z - x)$ by using the standard basis.

Problem 3 Let a linear transformation $T: \mathbb{R}^2 \to \mathbb{R}^2$ satisfy the following conditions: $T(1, 0) = (2, 3)$, $T(0, 1) = (-1, 1)$.

(1) Evaluate $T(-1, 1)$.

(2) Evaluate $T(x, y)$.

Problem 4 Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ moves any $\mathbf{x} \in \mathbb{R}^2$ to a symmetric image to a line which passes through the origin and has angle $\theta = \frac{\pi}{3}$ between the line and the x-axis. Find $T(\mathbf{x})$ for $\mathbf{x} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

Problem 5 Check whether the given matrix is an orthogonal matrix. If that is the case, find the inverse matrix.

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Solution $A^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ $\quad A A^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$\therefore$ A is an orthogonal matrix. And $A^{-1} = A^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$. ■

Problem 6 For each given linear transformation, find the kernel and range. Also determine whether it is bijective or not.

(1) $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 4x_1 - 2x_2 \\ x_2 - x_1 \end{bmatrix}$

(2) $S\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 + 9x_2 \\ -3x_2 - x_1 \end{bmatrix}$

Problem 7 Let T_1 and T_2 are defined as follows:

$T_1(x_1, x_2, x_3) = (4x_1, -2x_1 + x_2, -x_1 - 3x_2)$,

$T_2(x_1, x_2, x_3) = (x_1 + 2x_2, -x_3, 4x_1 - x_3)$.

(1) Find the standard matrix for each T_1 and T_2.

(2) Find the standard matrix for each $T_2 \circ T_1$ and $T_1 \circ T_2$.

Problem 8 Let $\mathbf{x}, \mathbf{z} \in \mathbb{R}^2$ be moved by two linear transformations T and S, where

$T(\mathbf{x}) = \begin{bmatrix} x_1 + 2x_2 \\ x_2 \end{bmatrix}$, $S(\mathbf{z}) = \begin{bmatrix} z_1 \\ -z_1 + z_2 \end{bmatrix}$.

Find $(S \circ T)(\mathbf{x})$.

Solution $[T] = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, $[S] = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ => $[S \circ T] = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$. ■

Problem 9 Answer the following questions.

(1) Find the dimension of the null space of the following matrix by using the Sage.

$$A=\begin{bmatrix} 2 & 5 & -3 & 7 & 1 & 3 \\ 5 & -2 & 9 & 8 & 4 & -2 \\ -4 & 3 & 8 & 11 & -5 & 2 \\ 11 & 0 & -2 & 4 & 10 & -1 \end{bmatrix}$$

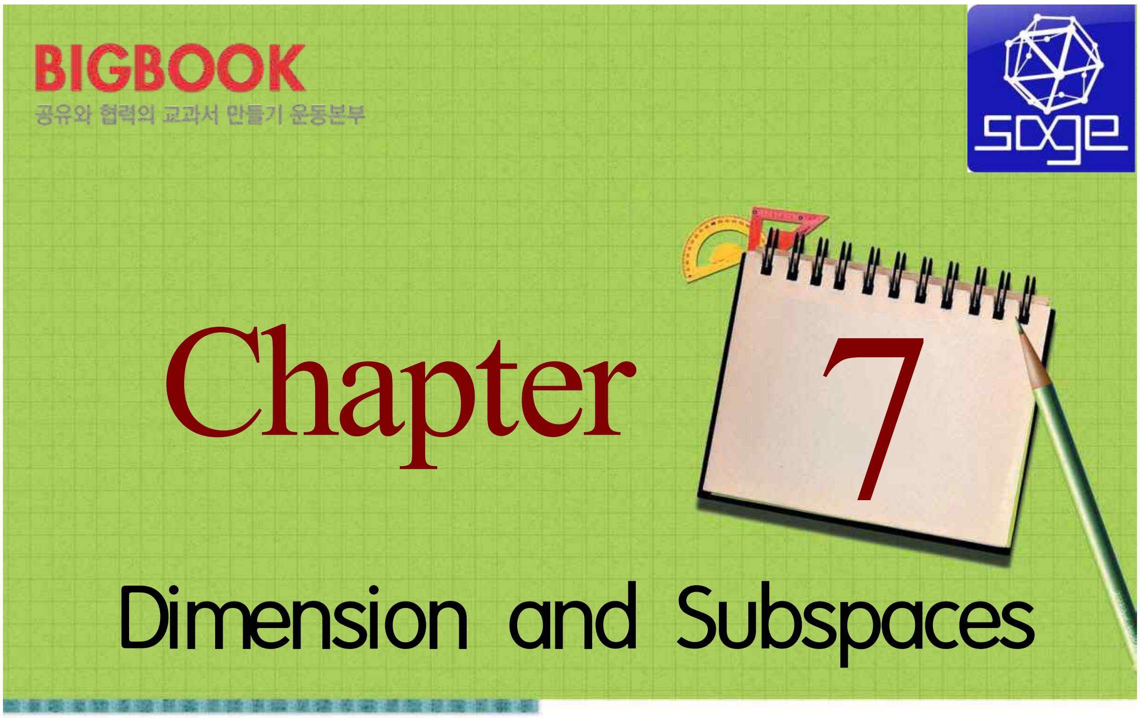
(2) Let T_A be a linear transformation corresponding to the above matrix A. Determine whether w$=(5, -2, -3, 6)$ is in the range of T_A by using the Sage.

Problem 10 Let $[R_\theta]=\begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ 0 & 0 & 1 \end{bmatrix}$. Find $[R_\theta]\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$ by using the Sage.

Solution

```
var('t')
var('x0')
var('y0')
A=matrix(3,3,[1, 0, x0, 0, 1, y0, 0, 0, 1]);
B=matrix(3,3,[cos(t), -sin(t), 0, sin(t), cos(t), 0, 0, 0, 1]);
C=matrix(3,3,[1, 0, -x0, 0, 1, -y0, 0, 0, 1]);
D=A*B*C
print D
var('x')
var('y')
E=matrix(3,1,[x, y, 1]);
F=D*E
print F
```

```
[x*cos(t) - x0*cos(t) - y*sin(t) + y0*sin(t) + x0]
[x*sin(t) - x0*sin(t) + y*cos(t) - y0*cos(t) + y0]
[                                               1]
```



Chapter 7

Dimension and Subspaces

The vector space $\mathbb{R}^n$ has a basis, **and it is a key concept to understand the vector space.** In particular, a basis provides a tool to compare sizes of different vector spaces with infinitely many elements. By understanding the size and structure of a vector space, one can visualize the space and efficiently use the data sitting contained within it.

In this chapter, we discuss bases and dimensions of vector spaces and then study their properties. We also study fundamental vector spaces associated with a matrix such as row space, column space, and nullspace, along with their properties. We then derive the Dimension Theorem describing the relationship between the dimensions of those spaces. In addition, the orthogonal projection of vectors in $\mathbb{R}^3$

will be generalized to vectors in $\mathbb{R}^n$, and we will study a standard matrix associated with an orthogonal projection which is a linear transformation. This matrix representation of an orthogonal projection will be used to study **Gram-Schmidt Orthogonalization** and **QR-Factorization**.

It will be shown that there are many different bases for $\mathbb{R}^n$, but the number of elements in every basis for $\mathbb{R}^n$ is always n. We also show that every nontrivial subspace of $\mathbb{R}^n$ has a basis, and study how to compute an orthogonal basis from the basis. Furthermore, we show how to represent a vector as a coordinate vector relative to a basis, which is not necessarily a standard basis, and find a matrix that maps a coordinate vector relative to a basis to a coordinator vector relative to another basis.

[Mathematicians in a Dish]

7.1 Properties of bases and dimensions

- Lecture Movie : http://youtu.be/or9c97J3Uk0, http://youtu.be/172stJmormk
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-9-sec-7-1.html

Having learned about standard bases, we will now discuss the concept of dimension of a vector space. Previously, we learned that an axis representing time can be added to the 3-dimensional physical space. We will now study the mathematical meaning of dimension. In this section, we define a basis and dimension of $\mathbb{R}^n$ using the concept of linear independence and study their properties.

Basis of a vector space

Definition [Basis]

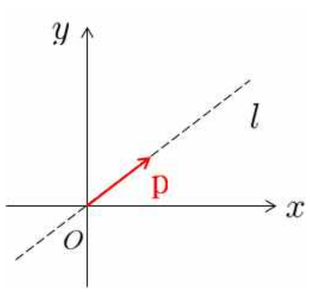
If a subset $S=\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s\}$ of $\mathbb{R}^n$ satisfies the following two conditions, then S is called a basis for $\mathbb{R}^n$:

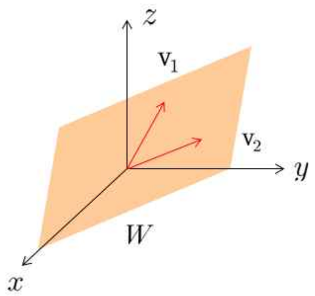
(1) S is linearly independent; and
(2) $\text{span}(S)=\mathbb{R}^n$.

Example 1

(1) If V is the subset of $\mathbb{R}^n$ consisting of all the points on a line going through the origin, then any nonzero vector in V forms a basis for V.

(2) If a subset V of $\mathbb{R}^n$ represents a plane going through the origin, then any two nonzero vectors in V that are not a scalar multiple of the other form a basis for V. ■





Example **2** Let $S=\{\mathbf{e}_1,\ \mathbf{e}_2\}$ where $\mathbf{e}_1=(1,0)$, $\mathbf{e}_2=(0,1)$. Since S is linearly independent and spans $\mathbb{R}^2$, S is a basis for $\mathbb{R}^2$. ■

- In general $S=\{\mathbf{e}_1,\ \mathbf{e}_2,\ \dots\ ,\ \mathbf{e}_n\}$ is a basis for $\mathbb{R}^n$, and it is called the **standard basis** for $\mathbb{R}^n$.

How to show linear independence of vectors in $\mathbb{R}^n$?

- Set of vectors $\mathbf{x}_1,\ \dots,\mathbf{x}_n$ in $\mathbb{R}^n$ is linear independent if

$$c_1\mathbf{x}_1+c_2\mathbf{x}_2+\cdots+c_m\mathbf{x}_m=\mathbf{0}\quad\Rightarrow\quad c_1=c_2=\cdots=c_m=0$$

- Let $A=[\mathbf{x}_1:\mathbf{x}_2:\ \cdots\ :\mathbf{x}_m]$ where $\mathbf{x}_i$'s are column vectors and $\mathbf{c}=[c_1\cdots c_m]^T$. If the homogeneous linear system $A\mathbf{c}=\mathbf{0}$ has the unique solution $\mathbf{c}=\mathbf{0}$, then the columns of the matrix A are linearly independent. In particular, for $m=n$, $\det A\neq 0$ implies the linear independence of the columns of A.

Theorem 7.1.1

The following n vectors in $\mathbb{R}^n$

$$\mathbf{x}_1=(x_{11},\ x_{12},\ \dots,\ x_{1n}),\ \dots,\ \mathbf{x}_n=(x_{n1},\ x_{n2},\ \dots,\ x_{nn})$$

are linearly independent if and only if

$$\Delta=\begin{vmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ \vdots & & \ddots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{nn}\end{vmatrix}\neq 0.$$

Proof For $c_1,\dots,c_n\in\mathbb{R}$,

$c_1\mathbf{x}_1+\ \cdots\ +c_n\mathbf{x}_n=\mathbf{0}$

$$\Rightarrow \; c_1\begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix} + c_2\begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix} + \cdots + c_n\begin{bmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{nn} \end{bmatrix} = \begin{bmatrix} c_1x_{11} + c_2x_{12} + \cdots + c_nx_{1n} \\ c_1x_{21} + c_2x_{22} + \cdots + c_nx_{2n} \\ \vdots \\ c_1x_{n1} + c_2x_{n2} + \cdots + c_nx_{nn} \end{bmatrix} = \mathbf{0}.$$

This gives us the following linear system

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

This linear system has the trivial solution $(c_1, c_2, \ldots, c_n) = \mathbf{0}$, i.e., $c_1 = 0, \ldots, c_n = 0$ if and only if $\Delta \neq 0$. Therefore $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly independent if and only if $\Delta \neq 0$. ■

Example 3 By Theorem 7.1.1, the following three vectors in $\mathbb{R}^3$

$$\mathbf{x}_1 = (1, 2, 3), \quad \mathbf{x}_2 = (-1, 0, 2), \quad \mathbf{x}_3 = (3, 1, 1)$$

are linearly independent because $\Delta = \begin{vmatrix} 1 & -1 & 3 \\ 2 & 0 & 1 \\ 3 & 2 & 1 \end{vmatrix} = 9 \neq 0$. □

- http://matrix.skku.ac.kr/RPG_English/7-TF-linearly-independent.html

Sage풀이 http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

```
x1=vector([1, 2, 3])
x2=vector([-1, 0, 2])
x3=vector([3, 1, 1])
A=column_matrix([x1, x2, x3])          # Generating the matrix with x1, x2,
                                       # x3 as its columns in that order
print A.det()
```

```
9
```
■

We can also use the inbuilt function of Sage to check if a set of vectors are linearly independent.

```
V=RR^3;x1=vector([1, 2, 3]);x2=vector([-1, 0, 2]);x3=vector([3, 1, 1])
S=[x1, x2, x3]
V.linear_dependence(S)
```

```
[]
```

Example 4 Show that $S=\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ with $\mathbf{x}_1=(1, 0, 0)$, $\mathbf{x}_2=(1, 1, 0)$, $\mathbf{x}_3=(1, 1, 1)$ is a basis for $\mathbb{R}^3$.

Solution

To show that $S=\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is a basis for $\mathbb{R}^3$, we need to show that S is linearly independent and it spans $\mathbb{R}^3$. □

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

```
A=matrix(QQ, 3, 3, [1, 1, 1, 0, 1, 1, 0, 0, 1])
print A.det()
```

```
1
```

■

- Since the computed determinant above is not zero, $S=\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is linearly independent. We now show that S spans $\mathbb{R}^3$. Let $\mathbf{x}=(x, y, z)$ be a vector in $\mathbb{R}^3$. Consider a linear system $\mathbf{x}=c_1\mathbf{x}_1+c_2\mathbf{x}_2+c_3\mathbf{x}_3$ in c_1, c_2, c_3. Note that if this linear system has a solution, then $\mathbf{x}=(x, y, z)$ is spanned by S. The linear system can be written as

$$\begin{aligned}(x, y, z) &= c_1(1, 0, 0)+c_2(1, 1, 0)+c_3(1, 1, 1),\ (c_i\in\mathbb{R})\\ &=(c_1+c_2+c_3,\ c_2+c_3,\ c_3),\end{aligned}$$

more explicitly, we have a linear system in c_1, c_2, c_3,

$$\begin{aligned} c_1+c_2+c_3&=x\\ c_2+c_3&=y\\ c_3&=z\end{aligned} \tag{1}$$

Hence we need to show that the linear system (1) has a solution to show that S spans R^3. Indeed, the coefficient matrix $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ of the linear system (1) is invertible, so the linear system (1) has a solution.

Theorem 7.1.2

Let $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be a basis for $\mathbb{R}^n$. For $r > n$, any subset $T = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_r\}$ of $\mathbb{R}^n$ is linearly dependent. Therefore, if T is linearly independent, then r must be less than or equal to n.

Proof http://matrix.skku.ac.kr/CLAMC/chap7/Page6.htm

Since S is a basis for $\mathbb{R}^n$, each vector in $T = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_r\}$ can be written as a linear combination of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. That is, there are $a_{ij} \in \mathbb{R}$ such that

$$\mathbf{y}_j = a_{1j}\mathbf{x}_1 + a_{2j}\mathbf{x}_2 + \cdots + a_{nj}\mathbf{x}_n = \sum_{i=1}^{n} a_{ij}\mathbf{x}_i, \quad (j = 1, 2, \ldots, r) \tag{2}$$

We now consider a formal equation with $c_1, c_2, \ldots, c_r \in \mathbb{R}$:

$$\sum_{j=1}^{r} c_j \mathbf{y}_j = c_1\mathbf{y}_1 + c_2\mathbf{y}_2 + \cdots + c_r\mathbf{y}_r = \mathbf{0}.$$

Then, from (2), we get,

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{r} a_{ij}\, c_j\right)\mathbf{x}_i = \sum_{j=1}^{r} c_j\left(\sum_{i=1}^{n} a_{ij}\mathbf{x}_i\right) = \mathbf{0}$$

Since $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are linearly independent,

$$\sum_{j=1}^{r} a_{ij}\, c_j = 0 \qquad (\forall\, i = 1, 2, \ldots, n)$$

Hence we get the following linear system

$$\begin{aligned} a_{11}c_1 + \cdots + a_{1r}c_r &= 0 \\ a_{21}c_1 + \cdots + a_{2r}c_r &= 0 \\ \vdots \qquad\quad & \vdots \\ a_{n1}c_1 + \cdots + a_{nr}c_r &= 0 \end{aligned} \tag{3}$$

The homogeneous linear system (3) has r unknowns, $c_1, c_2, \cdots, c_r$, and n linear equations. Since $r > n$, the linear system (3) must have a nontrivial solution. Therefore, T is linearly dependent. ■

Theorem 7.1.3

If $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ and $T = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_m\}$ are bases for $\mathbb{R}^n$, then $n = m$.

The proof of this theorem follows the theorem 7.1.2.

- There are infinitely many bases for $\mathbb{R}^n$. However, all the bases have the same number of vectors.

Definition [Dimension]

If S is a basis for $\mathbb{R}^n$, then the number of vectors in S is called the dimension of $\mathbb{R}^n$ and is denoted by $\dim \mathbb{R}^n$.

- Note that $\dim \mathbb{R}^n = n$. If its subspace V is the trivial subspace, $\{\mathbf{0}\}$, then $\dim V = 0$.

Theorem 7.1.4

For $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\} \subseteq \mathbb{R}^n$, the following holds:

(1) If S is linearly independent, then S is a basis for $\mathbb{R}^n$.

(2) If S spans $\mathbb{R}^n$ (i.e., $< S > = \mathbb{R}^n$), then S is a basis for $\mathbb{R}^n$.

Example 5 The determinant of the matrix having the vectors $\mathbf{x}_1 = (1, 2, -1)$, $\mathbf{x}_2 = (1, 3, 1)$, $\mathbf{x}_3 = (2, 0, 0)$ in $\mathbb{R}^3$ as its column vectors is

$$\begin{vmatrix} 1 & 1 & 2 \\ 2 & 3 & 0 \\ -1 & 1 & 0 \end{vmatrix} = 10 \neq 0.$$

Hence $S = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is linearly independent.

By Theorem 7.1.4, S is a basis for $\mathbb{R}^3$. ■

Theorem 7.1.5

If $S=\{\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_k\}$ is a basis for a subspace V of $\mathbb{R}^n$, then every vector v in V can be written as a unique linear combination of the vectors in S.

Proof Since S spans V, a vector v in V can be written as a linear combination of the vectors in S. Suppose

$$\mathrm{v}= t_1\mathrm{v}_1 + t_2\mathrm{v}_2 + \cdots + t_k\mathrm{v}_k \text{ and } \mathrm{v}= t'_1\mathrm{v}_1 + t'_2\mathrm{v}_2 + \cdots + t'_k\mathrm{v}_k.$$

By subtracting the second equation from the first one, we get

$$\mathbf{0}= (t_1 - t'_1)\mathrm{v}_1 + (t_2 - t'_2)\mathrm{v}_2 + \cdots + (t_k - t'_k)\mathrm{v}_k.$$

Since S is linearly independent, $t_1 - t'_1 = 0,\ t_2 - t'_2 = 0,\ \cdots,\ t_k - t'_k = 0$. Therefore $\mathrm{v}= t_1\mathrm{v}_1 + t_2\mathrm{v}_2 + \cdots + t_k\mathrm{v}_k$ is unique. ■

[Remark] Many a times a basis of $\mathbb{R}^n$ is defined to a set which satisfies conditions of theorem 7.1.5.

Example 6

Let $S=\{\mathrm{v}_1 = (1, 0, 0), \mathrm{v}_2 = (0, 1, 0), \mathrm{v}_3 = (0, 0, 1), \mathrm{v}_4 = (1, 1, 1)\}$. Then

$$\mathrm{v}= (3, 4, 5) = 3\mathrm{v}_1 + 4\mathrm{v}_2 + 5\mathrm{v}_3 + 0\mathrm{v}_4.$$

However, the vector v can also be written as follows:

$$\mathrm{v}= (3, 4, 5) = 4(1, 0, 0) + 5(0, 1, 0) + 6(0, 0, 1) - (1, 1, 1)$$

and

$$\mathrm{v}= (3, 4, 5) = 2(1, 0, 0) + 3(0, 1, 0) + 4(0, 0, 1) + (1, 1, 1).$$

This is possible because S is not a basis for $\mathbb{R}^3$. ■

7.2 Basic spaces of matrix

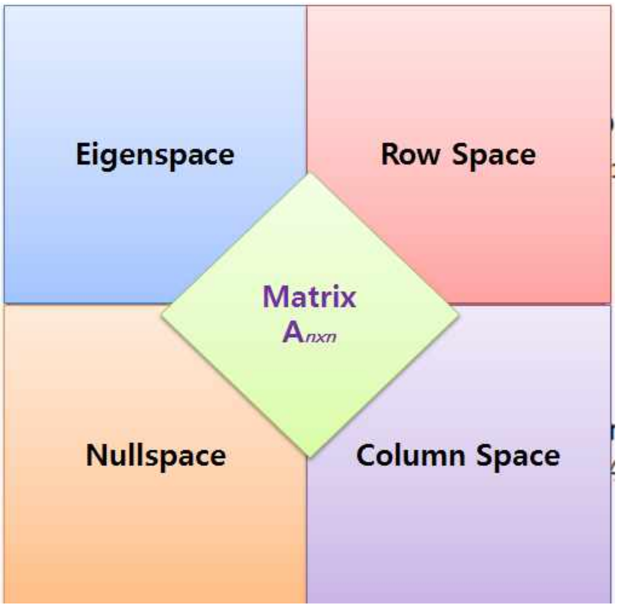
- Lecture Movie : http://youtu.be/KDM0-kBjRoM, http://youtu.be/8P7cd-Eh328
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-9-sec-7-2.html

Associated with an $m \times n$ matrix A, there are four important vector spaces: row space, column space, nullspace, and eigenspace. These vector spaces are crucial to study the algebraic and geometric properties of the matrix A as well as the solution space of a linear system having A as its coefficient matrix. In this section, we study the relationship between the column space and the row space of A and how to find a basis for the nullspace of A.

Eigenspace and null space

Definition [Solution space, Null space]

The eigenspace $\{x \in R^n \mid Ax = \lambda x\}$ of an $n \times n$ matrix A associated to an eigenvalue λ is a subspace of $\mathbb{R}^n$. The solution space of the homogeneous linear system $A\mathbf{x} = \mathbf{0}$ is also a subspace of $\mathbb{R}^n$. This is also called the **null space of** A and **denoted by** $\text{Null}(A)$.



Basis and dimension of a solution space

- Let A be an $n \times n$ matrix. For given augmented matrix $[A \vdots \ \mathbf{0}]$ of a homogeneous linear system with $A\mathbf{x}=\mathbf{0}$, by the Gauss-Jordan Elimination, we can get its RREF $[B \vdots \ \mathbf{0}]$. Suppose that matrix B has $r\,(1 \leq r \leq n)$ nonzero rows.

(1) If $r = n$, then the only solution to $A\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$. Hence the dimension of the solution space is zero.

(2) If $r < n$, then with permitting column exchanges, we can transform $[B \vdots \ \mathbf{0}]$ as

$$[B \vdots \ \mathbf{0}] \to \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & b_{1\,r+1} & \cdots & b_{1\,n} & \vdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & b_{2\,r+1} & \cdots & b_{2\,n} & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & b_{r\,r+1} & \cdots & b_{r\,n} & \vdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \vdots & 0 \end{bmatrix}.$$

Then the linear system is equivalent to

$$\begin{aligned} x_1 &= -b_{1\,r+1}x_{r+1} - b_{1\,r+2}x_{r+2} - \cdots - b_{1\,n}x_n \\ x_2 &= -b_{2\,r+1}x_{r+1} - b_{2\,r+2}x_{r+2} - \cdots - b_{2\,n}x_n \\ &\ \ \vdots \\ x_r &= -b_{r\,r+1}x_{r+1} - b_{r\,r+2}x_{r+2} - \cdots - b_{r\,n}x_n \end{aligned}$$

Here, $x_{r+1}, x_{r+2}, \ldots, x_n$ are $n-r$ free variables. Hence, for any real numbers $s_1, \ldots, s_{n-r}$, setting $x_{r+1} = s_1, \ldots, x_n = s_{n-r}$, any solution can be written as a linear combination of $n-r$ vectors as follows:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ x_{r+2} \\ x_{r+3} \\ \vdots \\ x_n \end{bmatrix} = s_1 \begin{bmatrix} -b_{1\,r+1} \\ -b_{2\,r+1} \\ \vdots \\ -b_{r\,r+1} \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + s_2 \begin{bmatrix} -b_{1\,r+2} \\ -b_{2\,r+2} \\ \vdots \\ -b_{r\,r+2} \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + s_{n-r} \begin{bmatrix} -b_{1\,n} \\ -b_{2\,n} \\ \vdots \\ -b_{r\,n} \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

Since $s_1, \ldots, s_{n-r}$ are arbitrary,

$$\mathbf{v}_1 = \begin{bmatrix} -b_{1\,r+1} \\ -b_{2\,r+1} \\ \vdots \\ -b_{r\,r+1} \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -b_{1\,r+2} \\ -b_{2\,r+2} \\ \vdots \\ -b_{r\,r+2} \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad \mathbf{v}_{n-r} = \begin{bmatrix} -b_{1\,n} \\ -b_{2\,n} \\ \vdots \\ -b_{r\,n} \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

are also solutions to the linear system. Hence, the previous linear combination of the $n-r$ vectors can be written as

$$\mathbf{x} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots + s_{n-r}\mathbf{v}_{n-r}.$$

This implies that $S=\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-r}\}$ spans the solution space of $A\mathbf{x}=\mathbf{0}$. In addition, it can be shown that S is linearly independent. Therefore S is a basis for the null space $\{\mathbf{x}\in\mathbb{R}^n \mid A\mathbf{x}=0\}$ of A and the dimension of the null space is $n-r$.

Definition **[Dimension of Null space]**

For an $m\times n$ matrix A, the dimension of the solution space of $A\mathbf{x}=\mathbf{0}$ is called the nullity of A and denoted by **nullity(A)**. That is, dim Null(A)=nullity(A).

Example 1

For the following matrix A, find a basis for the null space of A and the nullity of A.

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

Solution

The RREF of the augmented matrix $[A \,\vdots\, \mathbf{0}]$ for $A\mathbf{x}=\mathbf{0}$ is

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

Hence the general solution is

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -x_2 - 2x_4 \\ x_2 \\ x_4 \\ x_4 \end{bmatrix} = \begin{bmatrix} -s-2t \\ s \\ t \\ t \end{bmatrix} = s\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad (s,\ t \in \mathbb{R}).$$

Therefore a basis and the dimension of the null space of A is

$$S = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}, \text{ nullity}(A) = 2.$$

■

Example 2

Find a basis for the solution space of the following homogeneous linear system and its dimension.

$$\begin{aligned} 4x_1 + 12x_2 - 7x_3 + 6x_4 &= 0 \\ x_1 + 3x_2 - 2x_3 + x_4 &= 0 \\ 3x_1 + 9x_2 - 2x_3 + 11x_4 &= 0 \end{aligned}$$

Solution

Using Sage we can find the RREF of the coefficient matrix A:

```
A=matrix(ZZ, 3, 4, [4, 12, -7, 6, 1, 3, -2, 1, 3, 9, -2, 11])
print A.echelon_form()
```

```
[1 3 0 5]
[0 0 1 2]
[0 0 0 0]
```

Hence the linear system is equivalent to

$$\begin{aligned} x_1 &= -3x_2 - 5x_4 \\ x_3 &= \qquad\ \ -2x_4 \end{aligned}$$

Since x_2 and x_4 are free variables, letting $x_2 = r$, $x_4 = s$ for real numbers r, s, the solution can be written

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3r-5s \\ r \\ -2s \\ s \end{bmatrix} = r\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -5 \\ 0 \\ -2 \\ 1 \end{bmatrix}.$$

Hence we get the following basis and nullity:

$$S = \{(-3, 1, 0, 0),\ (-5, 0, -2, 1)\},\ \text{nullity}(A) = 2 \qquad \square$$

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

① Finding a basis for a null space

```
A=matrix(ZZ, 3, 4, [4, 12, -7, 6, 1, 3, -2, 1, 3, 9, -2, 11])
A.right_kernel()
```

```
Free module of degree 4 and rank 2 over Integer Ring
Echelon basis matrix:
[ 1  3  4 -2]
[ 0  5  6 -3]
```

② Computation of nullity

```
A.right_nullity()
```

```
2
```

■

Column space and row space

Definition

For given $m \times n$ matrix $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, the vectors obtained from the rows of A

$$A_{(1)} = [a_{11}\ \ a_{12}\ \ \cdots\ \ a_{1n}],\ A_{(2)} = [a_{21}\ \ a_{22}\ \ \cdots\ \ a_{2n}],$$

$$\ldots\ ,\ A_{(m)} = [a_{m1}\ \ a_{m2}\ \ \cdots\ \ a_{mn}]$$

are called **row vectors** and the vectors obtained from the columns of A

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

are called **column vectors**. The subspace of $\mathbb{R}^n$ spanned by the row vectors $A_{(1)}, \ldots, A_{(m)}$, that is,

$$< A_{(1)}, \ldots, A_{(m)} >$$

is called the **row space** of A and denoted by $\text{Row}(A)$. The subspace of $\mathbb{R}^m$ spanned by the column vectors $A^{(1)}, \ldots, A^{(n)}$, that is,

$$< A^{(1)}, \ldots, A^{(n)} >$$

is called the **column space** of A, and denoted by $\text{Col}(A)$. The dimension of the row space of A is called the **row rank of** A, and the dimension of the column space of A is called the **column rank of** A. The dimensions are denoted by $r(A)$ and $c(A)$, respectively, that is,

$$\dim\ \text{Row}(A) = r(A),\ \dim\ \text{Col}(A) = c(A)$$

Theorem 7.2.1

If two matrices A, B are row equivalent. then they have the same row space.

Proof

http://www.millersville.edu/~bikenaga/linear-algebra/matrix-subspaces/matrix-subspaces.html

- Note that the nonzero rows in the RREF of A form a basis for the row space of A. The same result can be applied to the column space of A.

Example 3

For the following set S, find a basis for $W = < S >$ which is a subspace of $\mathbb{R}^5$:

$$S = \{(1,2,1,3,2),\ (3,4,9,0,7),\ (2,3,5,1,8),\ (2,2,8,-3,5)\}$$

Solution

Note that the subspace W is equal to the row space of the following matrix

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

By Theorem **7.2.1**, it is also equal to the row space of the RREF of A

$$B = \begin{bmatrix} 1 & 0 & 7 & 0 & -39 \\ 0 & 1 & -3 & 0 & 31 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Therefore the collection of nonzero row vectors of B

$$\{(1,0,7,0,-39),\ (0,1,-3,0,31),\ (0,0,0,1,-7)\}$$

is a basis for $W = \mathrm{Row}(A)$.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

```
A=matrix(4, 5, [1, 2, 1, 3, 2, 3, 4, 9, 0, 7, 2, 3, 5, 1, 8, 2, 2, 8, -3, 5])
A.row_space()
```

```
Free module of degree 5 and rank 3 over Integer Ring
Echelon basis matrix:
[  1   0   7   0 -39]
[  0   1  -3   0  31]
[  0   0   0   1  -7]
```

■

Example 4

Find a basis for the column space of A:

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

Solution

The column space of A is equal to the row space of $A^T = \begin{bmatrix} 1 & 3 & 2 & 2 \\ 2 & 4 & 3 & 2 \\ 1 & 9 & 5 & 8 \\ 3 & 0 & 1 & -3 \\ 2 & 7 & 8 & 5 \end{bmatrix}$.

By Theorem **7.2.1**, it is also equal to the row space of the RREF of A^T:

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

Therefore $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ is a basis for the column space of A.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

```
A=matrix(4, 5, [1, 2, 1, 3, 2, 3, 4, 9, 0, 7, 2, 3, 5, 1, 8, 2, 2, 8, -3, 5])
A.column_space()
```

```
Free module of degree 4 and rank 3 over Integer Ring
Echelon basis matrix:
[ 1  0  0 -1]
[ 0  1  0  1]
[ 0  0  1  0]
```

■

Theorem 7.2.2

For $A = [a_{ij}]_{m \times n}$, the column rank and the row rank of A are equal.

For the proof of theorem 7.2.2, see http://mtts.org.in//expository-articles

- The same number for the column rank and the row rank of A is called the rank of A, and denoted by

$$r(A) = c(A) = \text{rank}(A)$$

[Remark] Relationship between vector spaces associated with a matrix A

- $\text{Row}(A^T) = \text{Col}(A)$, $\text{Col}(A^T) = \text{Row}(A)$,
- $\text{Row}(A)^\perp = \text{Null}(A)$, $\text{Null}(A)^\perp = \text{Row}(A)$,
- $\text{Col}(A)^\perp = \text{Null}(A^T)$, $\text{Null}(A^T)^\perp = \text{Col}(A)$

http://linear.ups.edu/html/section-CRS.html

Example 5 For $\mathbf{a}\,(\neq 0) \in \mathbb{R}^n$, $\mathbf{a}^\perp = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \mid a_1x_1 + \cdots + a_nx_n = 0\}$ is a hyperplane of $\mathbb{R}^n$. It is easy to see that $\mathbf{a}^\perp$ is a subspace of $\mathbb{R}^n$. ■

Example 6 (1) If $\mathbf{a} = (1, 2) \in \mathbb{R}^2$. Then

$$\mathbf{a}^\perp = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 + 2x_2 = 0\} = \{\alpha(-2, 1) \mid \alpha \in \mathbb{R}\}$$

is a line in the plane passing through the origin perpendicular to the vector $(1, 2)$.

(2) Let $\mathbf{a} = (1, 1, 1) \in \mathbb{R}^3$. Then

$$\mathbf{a}^\perp = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$$

is the plane in $\mathbb{R}^3$ passing through the origin and perpendicular to the vector $(1, 1, 1)$. ■

[International Linear Algebra Society] http://www.ilasic.org/

7.3 Dimension theorem (Rank-Nullity Theorem)

- Lecture Movie: http://youtu.be/ez7_JYRGsb4, http://youtu.be/bM-Pze0suqo
- Lab: http://matrix.skku.ac.kr/knou-knowls/cla-week-9-sec-7-3.html

In Section 7.2, we have studied the vector spaces associated to a matrix A. In this section, we study the relationship between the size of matrix A and the dimensions of the associated vector spaces.

Rank

Definition [rank]

The rank of a matrix A is defined to be the column rank (or the row rank) and denoted by $\text{rank}(A)$.

- Let A be an $m \times n$ matrix. If $U = \text{RREF}(A)$, then U can be written as the following:

$$U = \left.\left[\begin{array}{c:c} \multicolumn{2}{c}{\overbrace{\hphantom{I_r \quad\quad *}}^{n}} \\ I_r & * \\ \hdashline 0 & 0 \end{array}\right]\right\} m$$

Hence $\text{rank}(A) = r$ and $\text{nullity}(A) = n - r$.

Theorem 7.3.1 [Rank-Nullity theorem]

For any $A = [a_{ij}]_{m \times n}$, we have

$$\text{rank}(A) + \text{nullity}(A) = n$$

For the proof of theorem 7.3.1, see http://linear.ups.edu/html/section-IVLT.html

- The Rank-Nullity Theorem can be written as follows in terms of a linear transformation: If $A \in M_{m \times n}$ is the standard matrix for a linear transformation

$T : \mathbb{R}^n \rightarrow \mathbb{R}^m$, then

$$\dim(\mathrm{Im}(T)) = \mathrm{rank}(A),\ \dim(\ker(T)) = \mathrm{nullity}(A).$$

Hence

$$\dim(\mathrm{Im}(T)) + \dim(\ker(T)) = \dim(R^n) = n.$$

Example 1

The RREF of $A = \begin{bmatrix} 1 & 2 & 1 & 3 & 2 \\ 3 & 4 & 9 & 0 & 7 \\ 2 & 3 & 5 & 1 & 8 \\ 2 & 2 & 8 & -3 & 5 \end{bmatrix}$ is $B = \begin{bmatrix} 1 & 0 & 7 & 0 & -39 \\ 0 & 1 & -3 & 0 & 31 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. Hence rank($A$) $= 3$. Since $n = 5$, the dimension of the solution space for $A\mathbf{x} = \mathbf{0}$ is equal to nullity(A)$= 5 - 3 = 2$. ■

Example 2

Compute the rank and nullity of the matrix A, where

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

Solution

The RREF of A can be computed as follows

```
A = matrix(ZZ, 4, 5, [1, -2, 1, 1, 2, -1, 3, 0, 2, -1, 0, 1, 1, 3, 4, 1, 2, 5, 13, 5])
A.echelon_form()
```

```
[1 0 3 7 0]
[0 1 1 3 0]
[0 0 0 0 1]
[0 0 0 0 0]
```

Hence rank(A)$= 3$, and by Theorem **7.3.1**, $\mathrm{nullity}(A) = 5 - \mathrm{rank}(A) = 5 - 3 = 2$. ■

- http://matrix.skku.ac.kr/RPG_English/7-B2-rank-nullity.html

Sage http://sage.skku.edu

```
print A.rank()                # rank computation
print A.right_nullity()       # nullity computation
```

```
3
2
```
■

Theorem 7.3.2

A linear system $A\mathbf{x} = \mathbf{b}$ has a solution if and only if

$$\mathrm{rank}(A) = \mathrm{rank}[A\ \vdots\ \mathbf{b}].$$

Proof Let $A = [a_{ij}]_{m\times n}$, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{b} = (b_1, b_2, \ldots, b_m)$. Then the linear system $A\mathbf{x} = \mathbf{b}$ can be written as

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}. \qquad (1)$$

Hence we have the following:

$A\mathbf{x} = \mathbf{b}$ has a solution. $\Leftrightarrow$ There exist $x_1, x_2, \ldots, x_n$ satisfying the linear system (1).

$\Leftrightarrow$ $\mathbf{b}$ is a linear combination of the columns of A.

$\Leftrightarrow$ $\mathbf{b} \in \mathrm{Col}(A)$

$\Leftrightarrow$ $\mathrm{rank}(A) = \mathrm{rank}[A\ \vdots\ \mathbf{b}]$. ■

Example 3

The linear system $\begin{aligned} x-2y+2z &= 1 \\ x+4y+3z &= 2 \\ 2x+2y+5z &= 3 \end{aligned}$ has its matrix-vector representation

$$\begin{bmatrix} 1 & -2 & 2 \\ 1 & 4 & 3 \\ 2 & 2 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

```
A = matrix(ZZ, 3, 3, [1, -2, 2, 1, 4, 3, 2, 2, 5])
b = vector([1, 2, 3])
print A.rank()                      # rank(A)
print A.augment(b).rank()           # rank[A : b]
```

```
2
2
```

Since $\text{rank}(A) = 2 = \text{rank}[A : \mathbf{b}]$, Theorem **7.3.2** implies that the linear system has a solution. ■

Definition **[Hyperplane]**

Let $\mathbf{a} \in \mathbb{R}^n$ be a nonzero vector. Then

$\mathbf{a}^{\perp} = \{\mathbf{x} \in R^n \mid \mathbf{a} \cdot \mathbf{x} = 0\}$ is called the orthogonal complement of $\mathbf{a}$.

(This can be understood as the solution space of $\mathbf{a} \cdot \mathbf{x} = \mathbf{x}^T \mathbf{a} = 0$.) The orthogonal complement of $\mathbf{a}$ is a hyperplane of $\mathbb{R}^n$.

- Note that $\dim \mathbf{a}^{\perp} = \text{nullity}(\mathbf{a}^T) = n-1$.

Theorem 7.3.3

Let W be a $n-1$ dimensional subspace of $\mathbb{R}^n$. Then $W = \mathbf{a}^{\perp}$ for some nonzero vector $\mathbf{a} \in R^n$.

Proof Since $\dim W = n-1$, by the Rank-Nullity Theorem, $\dim W^{\perp} = 1$. Thus $W^{\perp} = \text{span}\{\mathbf{a}\}$ for a nonzero vector $\mathbf{a}$. Therefore

$W = (W^{\perp})^{\perp} = (\text{span}\{\mathbf{a}\})^{\perp} = \mathbf{a}^{\perp}$. ■

7.4 Rank theorem

- Lecture Movie : http://youtu.be/8P7cd-Eh328 http://youtu.be/bM-Pze0suqo
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-9-sec-7-4.html

In this section, we study the relationship between the rank of a matrix A and the theorems that is related to the dimension of subspaces associated to A.

Theorem 7.4.1 [Rank theorem]

For any $A=[a_{ij}]_{m\times n}$, dim Row(A)=dim Col(A).

Proof

http://ocw.mit.edu/courses/mathematics/18-701-algebra-i-fall-2010/study-materials/MIT18_701F10_rrk_crk.pdf ■

Theorem 7.4.2

For any $A=[a_{ij}]_{m\times n}$, rank(A) $\leq$ min $\{m, n\}$.

Proof Since dim Row(A) $\leq m$, dim Col(A) $\leq n$, and rank(A)=dim Row(A)=dim Col(A), it follows that rank(A) $\leq$ min $\{m, n\}$ ■

Theorem 7.4.3 [Rank theorem]

Given $A=[a_{ij}]_{m\times n}$, the followings hold:

(1) dim Row(A)+dim Null(A)= the number of columns of A(that is, rank(A)+nullity(A)=n).

(2) dim Col(A)+dim Null(A^T)=the number of rows of A(that is, rank(A)+nullity(A^T)=m).

Proof (1) follows from Theorem **7.3.1**,

(2) follows from the fact that Row(A^T)=Col(A) and rank(A)=rank(A^T)

along with replacing A in (1) by A^T. ■

Theorem 7.4.4

For a square matrix A of order n, A is invertible if and only if

$$\text{rank}(A) = n.$$

Proof If A is invertible, then $A\mathbf{x} = \mathbf{0}$ has the trivial solution only and hence $\text{Null}(A) = \{\mathbf{0}\}$, giving $\text{nullity}(A) = 0$. By the Rank-Nullity Theorem, we have $\text{rank}(A) = n$. This can be reversed. ■

Example 1 Find the rank and nullity of the following matrix:

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

Solution

Using Gaussian Elimination,

$$\begin{bmatrix} 1 & 3 & 1 & 7 \\ 2 & 3 & -1 & 9 \\ -1 & -2 & 0 & -5 \end{bmatrix} \xrightarrow[R_1 + R_3]{(-2)R_1 + R_2} \begin{bmatrix} 1 & 3 & 1 & 7 \\ 0 & -3 & -3 & -5 \\ 0 & 1 & 1 & 2 \end{bmatrix}$$

$$\xrightarrow[(1/3)R_2 + R_3]{(-1)R_2} \begin{bmatrix} 1 & 3 & 1 & 7 \\ 0 & 3 & 3 & 5 \\ 0 & 0 & 0 & \frac{1}{3} \end{bmatrix}$$

$$\xrightarrow[3R_3]{(1/3)R_2} \begin{bmatrix} 1 & 3 & 1 & 7 \\ 0 & 1 & 1 & \frac{5}{3} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{(-3)R_2 + R_1} \begin{bmatrix} 1 & 0 & -2 & 2 \\ 0 & 1 & 1 & \frac{5}{3} \\ 0 & 0 & 0 & 1 \end{bmatrix} = U = \text{REF}(A).$$

Hence $\text{rank}(A) = 3$ and the Rank-Nullity Theorem gives $4 - \text{rank}(A) = 4 - 3 = 1 = \text{nullity}(A)$.

Sage http://sage.skku.edu

```
A=matrix(3, 4, [1, 3, 1, 7, 2, 3, -1, 9, -1, -2, 0, -5])
print A.rank()                    # rank computation
```

```
print A.right_nullity()            # nullity computation
```

```
3
1
```

■

Theorem 7.4.5

For matrices A, B with multiplication AB defined, the followings hold:

(1) Null(B)⊆ Null(AB).

(2) Null(A^T)⊆ Null($(AB)^T$).

(3) Col(AB)⊆ Col(A).

(4) Row(AB)⊆ Row(B).

Proof We prove only (1) here. For

$\mathrm{x} \in \mathrm{Null}(B) \Rightarrow B\mathrm{x} = \mathbf{0} \Rightarrow (AB)\mathrm{x} = A(B\mathrm{x}) = A\mathbf{0} = \mathbf{0}$.

$\therefore \mathrm{x} \in \mathrm{Null}(AB)$ ■

Theorem 7.4.6

rank(AB) ≤ min{rank(A), rank(B)}.

Follows from theorem 7.4.5.

Theorem 7.4.7

Multiplying a matrix B by an invertible matrix $A=[a_{ij}]_{n\times n}$ does not change the rank of B. That is, if $|A| \neq 0$, then

$$\mathrm{rank}(AB) = \mathrm{rank}(B) = \mathrm{rank}(BA).$$

Follows from theorem 7.4.6.

Theorem 7.4.8

Suppose $A=[a_{ij}]_{m\times n}$ has rank(A)= r. Then

(1) Every submatrix C of A satisfies rank(C) ≤ r.
(2) A must have at least one $r\times r$ submatrix whose rank is equal to r.

Proof (1) Suppose the submatrix C is obtained by taking s rows of A (we let B be this matrix consisting of the s rows of A) and taking t columns of B. Since Row(B) ⊆ Row(A) and Col(C) ⊆ Col(B), the result follows.

(2) Since the rank of A is r, there are r linearly independent rows of A. Then the matrix B consisting of the r linearly independent rows has the rank equal to r. We now form a matrix C by taking r linearly independent columns of B. Then C is an $r\times r$ submatrix of A whose rank is equal to r. ■

Main Theorem of Inverse Matrices

Theorem 7.4.9 [Equivalent statements of invertible matrices]

For an $n\times n$ matrix A, the following are equivalent:
(1) A is invertible.
(2) $\det(A)\neq 0$.
(3) A is equivalent to I_n.
(4) A is a product of elementary matrices.
*(5) A has a unique LDU-factorization. That is, there exists a permutation matrix P such that $PA=LDU$ where L is a lower triangular matrix with all the diagonal entries equal to 1, D is an invertible diagonal matrix, and U is an upper triangular matrix whose main diagonal entries are all equal to 1.
(6) For any $n\times 1$ vector $\mathbf{b}$, $A\mathbf{x}=\mathbf{b}$ has a unique solution.
(7) $A\mathbf{x}=\mathbf{0}$ has the unique solution $\mathbf{x}=\mathbf{0}$.
(8) The column vectors of A are linearly independent.
(9) The column vectors of A span R^n.
*(10) A has a left inverse. That is, there exists a matrix X of order n such that $XA=I$.

(11) $\operatorname{rank}(A) = n$.

(12) The row vectors of A are linearly independent.

(13) The row vectors of A span R^n.

*(14) A has a right inverse. That is, there exists a matrix X of order n satisfying $AX = I$.

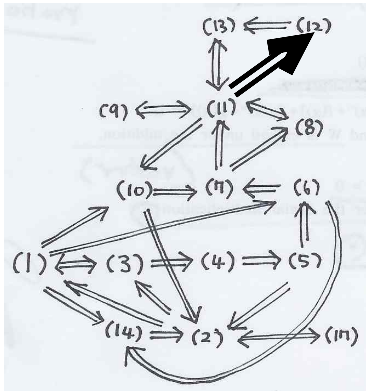
(15) T_A is one-to-one.

(16) T_A is onto.

(17) $\lambda = 0$ is not an eigenvalue of A.

(18) $\operatorname{nullity}(A) = 0$.

Proof We first prove the following equivalence:



① (10) ⇒ (7) ⇒ (8) ⇒ (11) ⇒ (10)

(10) ⇒ (7): Suppose A has a left inverse X such that $XA = I$. If $\mathbf{x}$ satisfies $A\mathbf{x} = \mathbf{0}$, then $XA = I$ gives

$$\mathbf{x} = I\mathbf{x} = (XA)\mathbf{x} = X(A\mathbf{x}) = X\mathbf{0} = \mathbf{0}.$$

Hence $A\mathbf{x} = \mathbf{0}$ has the unique solution $\mathbf{x} = \mathbf{0}$.

(7) ⇒ (8): Suppose $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. If $\mathbf{v}_k$ denotes the kth column vector of A and $\mathbf{x} = [\alpha_1\ \alpha_2\ \cdots\ \alpha_n]^T$, then

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots \alpha_n\mathbf{v}_n = \mathbf{0} \ \Leftrightarrow\ A\mathbf{x} = \mathbf{0} \ \Rightarrow\ \mathbf{x} = \mathbf{0} \ \Leftrightarrow\ \alpha_i = 0 \quad 1 \le i \le n$$

Hence the set $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ of the column vectors of A is linearly independent.

(8) ⇒ (11): Suppose the column vectors of A are linearly independent. Then $\operatorname{rank}(A)$, which is equal to the maximum number of linearly independent columns of A, is equal to n.

(11) ⇒ (10): Suppose $\operatorname{rank}(A) = n$. Then the rows of A are linearly independent. Let $\mathbf{e}_k$ be the kth standard basis vector. Then the following linear systems

$$A^T\mathbf{x} = \mathbf{e}_k, \quad 1 \le k \le n$$

are consistent for all k, since $\text{rank}(A^T) = n = \text{rank}\left[A^T : \mathbf{e}_k\right]$. Letting $\mathbf{x}_k$ be a solution to the linear systems, $X = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_n^T \end{bmatrix}$ is a left inverse of A.

② (1) ⇒ (6) ⇒ (14) ⇒ (2) ⇒ (1)

(1) ⇒ (6): Suppose A is invertible. Then, for any $n \times 1$ vector $\mathbf{b}$,

$$A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}.$$

Hence $A\mathbf{x} = \mathbf{b}$ has a solution $\mathbf{x}_0 = A^{-1}\mathbf{b}$. For the uniqueness of the solution, suppose $\mathbf{x}$ is another solution. Then

$$\mathbf{x} = I\mathbf{x} = (A^{-1}A)\mathbf{x} = A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{b} = \mathbf{x}_0.$$

Therefore $A\mathbf{x} = \mathbf{b}$ has a unique solution.
(6) ⇒ (14): Suppose that for each $n \times 1$ $\mathbf{b}$, the linear system $A\mathbf{x} = \mathbf{b}$ has a unique solution. If we take $\mathbf{b}$ to be $\mathbf{e}_k$, the kth standard basis vector, then the following linear system

$$A\mathbf{x} = \mathbf{e}_k, \quad 1 \le k \le n$$

also has a unique solution. If $\mathbf{x}_k$ is the solution to the linear system, then the matrix $X = \left[\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_n\right]$ is a right inverse of A.
(14) ⇒ (2): Suppose A has a right inverse X such that $AX = I$. Then

$$\det(A)\det(X) = \det(AX) = \det(I) = 1.$$

Hence $\det(A) \neq 0$.
(2) ⇒ (1): Suppose $\det(A) \neq 0$. If we let $B = \dfrac{1}{\det(A)} \text{adj}(A)$, then it can be shown that

$$AB = BA = I.$$

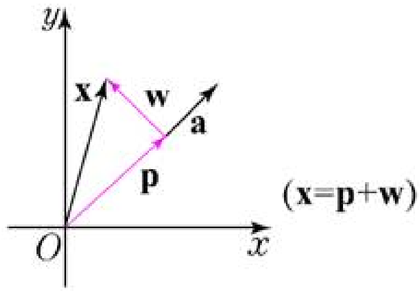
Hence A is invertible. ■

7.5 Projection Theorem

- Lecture Movie : http://youtu.be/GlcA4l8SmlM, http://youtu.be/Rv1rd3u-oYg
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-10-sec-7-5.html

In Chapter 1, we have studied the orthogonal project in $\mathbb{R}^3$ where the vectors and their projecttions can be visualized. In this section, we generalize the concept of project in $\mathbb{R}^n$. We also show that the projection is a linear transformation and find its standard matrix, which will be crucial to study the Gram-Schmidt Orthogonalization and the QR-Decomposition.

Orthogonal Projection in $\mathbb{R}^2$



$$\mathbf{p} = t\mathbf{a} = \frac{\mathbf{x}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} = \text{proj}_{\mathbf{a}}\mathbf{x}$$

$$\mathbf{w} = \mathbf{x} - \mathbf{p}$$

Projection (in 1-Dimension subspace) on $\mathbb{R}^n$

Theorem 7.5.1 [Projection]

For any nonzero vector $\mathbf{a}$ in $\mathbb{R}^n$, every vector $\mathbf{x} \in \mathbb{R}^n$ can be expressed as follows:

$$\mathbf{x} = \text{proj}_{<\mathbf{a}>}\mathbf{x} + \mathbf{w} = t\mathbf{a} + \mathbf{w} = \mathbf{p} + \mathbf{w},$$

where $\mathbf{p}$ is a scalar multiple of $\mathbf{a}$ and $\mathbf{w}$ is perpendicular to $\mathbf{a}$. Furthermore, the vectors $\mathbf{p}$, $\mathbf{w}$ can be written as follows:

$$\mathbf{p} = t\mathbf{a} = \frac{\mathbf{x}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a}, \quad \mathbf{w} = \mathbf{x} - \mathbf{p}.$$

The proof of the above theorem is similar to that in case of orthogonal projection in the $\mathbb{R}^2$ and $\mathbb{R}^3$.

- In the above theorem, the vector $\mathbf{p}$ is called the orthogonal projection of $\mathbf{x}$ onto $\text{span}\{\mathbf{a}\}$ and denoted by $\text{proj}_{<\mathbf{a}>}\mathbf{x} = \dfrac{\mathbf{x}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a}$. The vector $\mathbf{w}$ is called the orthogonal complement of the vector $\mathbf{a}$.

Definition **[Orthogonal projection on $\mathbb{R}^n$]**

The transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ defined below

$$T(\mathbf{x}) = \text{proj}_{<\mathbf{a}>}\mathbf{x} = \frac{\mathbf{x}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a}$$

is called the orthogonal projection of $\mathbb{R}^n$ onto $\text{span}\{\mathbf{a}\}$.

- It can be shown that the orthogonal projection $T(\mathbf{x}) = \text{proj}_{<\mathbf{a}>}\mathbf{x}$ is a linear transformation.
(http://www.math.lsa.umich.edu/~speyer/417/OrthoProj.pdf)

Theorem 7.5.2

Let $\mathbf{a}$ be a nonzero column vector in $\mathbb{R}^n$. Then the standard matrix of

$$T(\mathbf{x}) = \text{proj}_{<\mathbf{a}>}\mathbf{x} = P\mathbf{x}$$

is

$$P = \frac{1}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T.$$

Note that P is a symmetric matrix and $\text{rank}(P) = 1$.

For the proof of this theorem, see the website:
http://ocw.mit.edu/courses/mathematics/18-06sc-linear-algebra-fall-2011/least-squares-determinants-and-eigenvalues/projections-onto-subspaces/MIT18_06SCF11_Ses2.2sum.pdf

Example 1

Using the above theorem, find the standard matrix P_θ of the orthogonal projection in $\mathbb{R}^2$ onto the line $y=(\tan\theta)x$ passing through the origin.

Solution

This is a problem of finding the orthogonal projection of a vector $\mathbf{x}$ onto the subspace spanned by a vector $\mathbf{a}$. Hence we take $\mathbf{a}$ as a unit vector $\mathbf{u}$ on the line $y=(\tan\theta)x$. Since the slope of the line is $\tan\theta=\dfrac{\sin\theta}{\cos\theta}$, $\mathbf{u}=\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}$ and $\|\mathbf{u}\|^2=1$. Therefore, by the previous theorem,

$$P_\theta=\frac{1}{\mathbf{u}^T\mathbf{u}}\mathbf{u}\mathbf{u}^T=\frac{1}{\|\mathbf{u}\|^2}\mathbf{u}\mathbf{u}^T=\mathbf{u}\mathbf{u}^T=\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}[\cos\theta\ \sin\theta]=\begin{bmatrix}\cos^2\theta & \sin\theta\cos\theta\\ \sin\theta\cos\theta & \sin^2\theta\end{bmatrix}\ ■$$

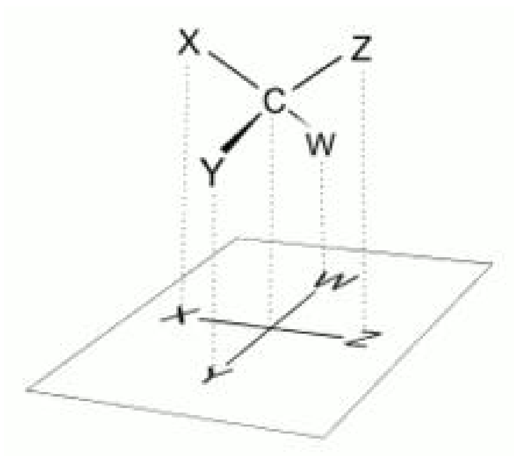
Example 2

Find the standard matrix P for the orthogonal projection T in $\mathbb{R}^3$ onto the subspace spanned by the vector $\mathbf{a}=(1,-4,2)$.

Solution

$$\mathbf{a}^T\mathbf{a}=[1\ -4\ \ 2]\begin{bmatrix}1\\-4\\2\end{bmatrix}=21,\quad \mathbf{a}\mathbf{a}^T=\begin{bmatrix}1\\-4\\2\end{bmatrix}[1\ -4\ 2]=\begin{bmatrix}1&-4&2\\-4&16&-8\\2&-8&4\end{bmatrix}$$

Hence, $P=\dfrac{1}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T=\dfrac{1}{21}\begin{bmatrix}1&-4&2\\-4&16&-8\\2&-8&4\end{bmatrix}$ ■



WIKIMEDIA COMMONS

http://en.wikipedia.org/wiki/Fischer_projection

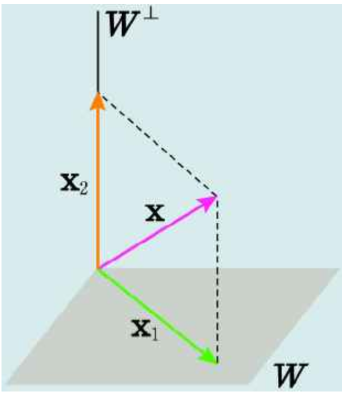
Projection of $\mathbf{x}$ on subspace W in $\mathbb{R}^n$

Theorem 7.5.3

Let W be a subspace of $\mathbb{R}^n$. Then every vector $\mathbf{x}$ in $\mathbb{R}^n$ can be uniquely expressed as follows:

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 \quad \text{where } \mathbf{x}_1 \in W \text{ and } \mathbf{x}_2 \in W^{\perp}.$$

In this case $\mathbf{x}_1$ is called the orthogonal projection of $\mathbf{x}$ onto W and is denoted by $\mathrm{proj}_W \mathbf{x}$.



$$\mathbf{x}_1 = \mathrm{proj}_W \mathbf{x}, \quad \mathbf{x}_2 = \mathbf{x} - \mathbf{x}_1 = \mathrm{proj}_{W^{\perp}} \mathbf{x}$$

http://www.math.lsa.umich.edu/~speyer/417/OrthoProj.pdf

Theorem 7.5.4

Let W be a subspace of $\mathbb{R}^n$. If M is a matrix whose columns are the vectors in a basis for W, then for each vector $\mathbf{x} \in \mathbb{R}^n$

$$\mathrm{proj}_W \mathbf{x} = M(M^T M)^{-1} M^T \mathbf{x}.$$

Proof http://www.math.lsa.umich.edu/~speyer/417/OrthoProj.pdf

Example 3 Find the standard matrix for the orthogonal projection in $\mathbb{R}^3$ onto the plane $x-4y+2z=0$.

Solution

The general solution to $x-4y+2z=0$ is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4t_1-2t_2 \\ t_1 \\ t_2 \end{bmatrix} = t_1\begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix} + t_2\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \quad (t_1,\ t_2 \in \mathbb{R}).$$

Thus $\{(4,1,0),(-2,0,1)\}$ is a basis of the plane $x-4y+2z=0$.

Hence, by taking $M=\begin{bmatrix} 4 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$, the standard matrix is $P=M(M^TM)^{-1}M^T$.

Since $M^TM=\begin{bmatrix} 4 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}\begin{bmatrix} 4 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 17 & -8 \\ -8 & 5 \end{bmatrix}$ and $(M^TM)^{-1} = \begin{bmatrix} \frac{5}{21} & \frac{8}{21} \\ \frac{8}{21} & \frac{17}{21} \end{bmatrix}$,

$$P=M(M^TM)^{-1}M^T=\begin{bmatrix} 4 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \frac{5}{21} & \frac{8}{21} \\ \frac{8}{21} & \frac{17}{21} \end{bmatrix}\begin{bmatrix} 4 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{20}{21} & \frac{4}{21} & -\frac{2}{21} \\ \frac{4}{21} & \frac{5}{21} & \frac{8}{21} \\ -\frac{2}{21} & \frac{8}{21} & \frac{17}{21} \end{bmatrix}$$

Sage http://sage.skku.edu

```
M=matrix(3, 2, [4, -2, 1, 0, 0, 1])
print M*(M.transpose()*M).inverse()*M.transpose()
```

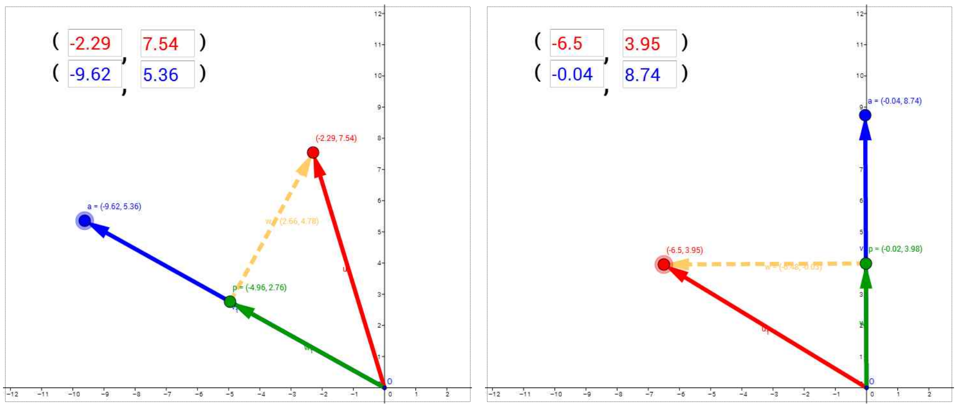
```
[20/21   4/21  -2/21]
[ 4/21   5/21   8/21]
[-2/21   8/21  17/21]
```

■

- The standard matrix P for an orthogonal projection is symmetric and idempotent ($P^2=P$).

[Remark] Simulation of the projection of two vectors

- http://www.geogebratube.org/student/m9503



http://matrix.skku.ac.kr/mathLib/main.html

7.6 * Least square solutions

- Lecture Movie : https://youtu.be/BC9qeR0JWis
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-10-sec-7-6.html

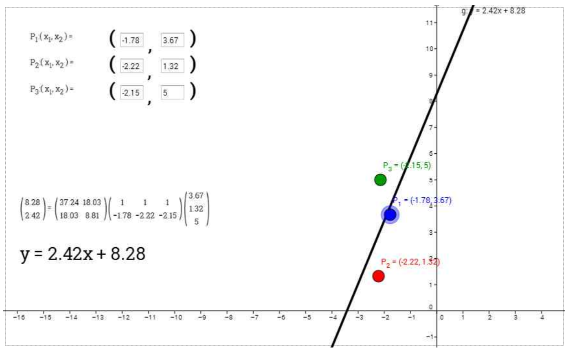
Previously, we have studied how to find solve the linear system $A\mathbf{x}=\mathbf{b}$ when the linear system has a solution. In this section, we study how to find an optimal solution using projection when the linear system does not have any solution.

Details can be found in the following websites:

http://www.seas.ucla.edu/~vandenbe/103/lectures/ls.pdf

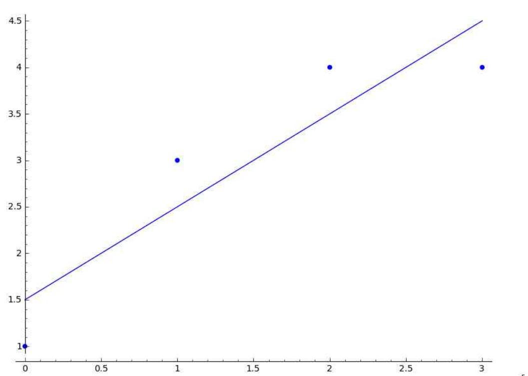
Least square solutions with GeoGebra

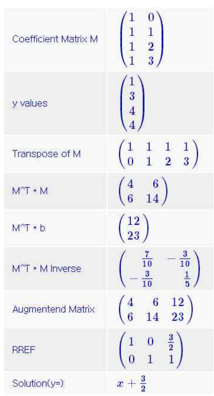
<Simulations> http://www.geogebratube.org/student/m12933



Least square solutions with Sage

<Simulations> http://matrix.skku.ac.kr/2012-album/11.html





Coefficient Matrix M	$\begin{pmatrix}1&0\\1&1\\1&2\\1&3\end{pmatrix}$
y values	$\begin{pmatrix}1\\3\\4\\4\end{pmatrix}$
Transpose of M	$\begin{pmatrix}1&1&1&1\\0&1&2&3\end{pmatrix}$
M^T * M	$\begin{pmatrix}4&6\\6&14\end{pmatrix}$
M^T * b	$\begin{pmatrix}12\\23\end{pmatrix}$
M^T * M Inverse	$\begin{pmatrix}\frac{7}{10}&-\frac{3}{10}\\-\frac{3}{10}&\frac{1}{5}\end{pmatrix}$
Augmentend Matrix	$\begin{pmatrix}4&6&12\\6&14&23\end{pmatrix}$
RREF	$\begin{pmatrix}1&0&\frac{3}{2}\\0&1&1\end{pmatrix}$
Solution(y=)	$x+\frac{3}{2}$

7.7 Gram-Schmidt Orthonomalization process

- Lecture Movie: http://youtu.be/gt4-EuXvx1Y, http://youtu.be/EBCi1nR7EuE
- Lab: http://matrix.skku.ac.kr/knou-knowls/cla-week-10-sec-7-7.html

Every basis of $\mathbb{R}^n$ has n elements, but all the bases are distinct. In this section, we show that every nontrivial subspace of $\mathbb{R}^n$ has a basis and how to find an orthonormal basis from a given basis.

[Remark]

The subspaces $\{\mathbf{0}\}$ and $\mathbb{R}^n$ of $\mathbb{R}^n$ are called trivial subspaces.
There are many different bases for $\mathbb{R}^n$, but all the bases have n elements and the number n is called the dimension of $\mathbb{R}^n$.

Orthogonal set and orthonormal set

Definition

For vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ in $\mathbb{R}^n$, let

$$S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}.$$

If every pair of vectors in S is orthogonal, then S is called an orthogonal set. Furthermore, if every vector in the orthogonal set S is a unit vector, then S is called an orthonormal set.

- The above definition can be summarized as follows:

S is an orthogonal set. $\Leftrightarrow$ $\mathbf{x}_i \cdot \mathbf{x}_j = 0 \quad (i \neq j)$

S is an orthonormal set. $\Leftrightarrow$ $\mathrm{x}_i \cdot \mathrm{x}_j = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases}$ $(=\delta_{ij}$ Kronecker delta)

Example 1

(1) The standard basis $\{\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3\}$ for R^3 is orthonormal.

(2) In R^3 let $\mathrm{x}_1 = (0,1,0)$, $\mathrm{x}_2 = (2,0,1)$, $\mathrm{x}_3 = (1,0,-2)$. Then $\{\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3\}$ is orthogonal, but not orthonormal.

(3) In R^3 let $\mathrm{y}_1 = (0,1,0)$, $\mathrm{y}_2 = \left(\frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}}\right)$, $\mathrm{y}_3 = \left(\frac{1}{\sqrt{5}}, 0, -\frac{2}{\sqrt{5}}\right)$. Then the set $\{\mathrm{y}_1, \mathrm{y}_2, \mathrm{y}_3\}$ is orthonormal.

(4) If $\{\mathrm{x}_1, \ldots, \mathrm{x}_n\}$ is an orthogonal set, then $\left\{\frac{\mathrm{x}_1}{\|\mathrm{x}_1\|}, \ldots, \frac{\mathrm{x}_n}{\|\mathrm{x}_n\|}\right\}$ is an orthonormal set. ■

Orthogonality and Linear independance

Theorem 7.7.1

Let $S = \{\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_k\}$ be a set of nonzero vectors in $\mathbb{R}^n$. If S is orthogonal, then S is linearly independent.

Proof For $c_1, c_2, \ldots, c_k \in \mathbb{R}$, suppose

$$c_1\mathrm{x}_1 + c_2\mathrm{x}_2 + \cdots + c_k\mathrm{x}_k = \mathbf{0}.$$

Then, for each i $(i = 1, 2, \ldots, k)$,

$$(c_1\mathrm{x}_1 + c_2\mathrm{x}_2 + \cdots + c_k\mathrm{x}_k) \cdot \mathrm{x}_i = \mathbf{0} \cdot \mathrm{x}_i.$$

That is,

$$c_1(\mathrm{x}_1 \cdot \mathrm{x}_i) + c_2(\mathrm{x}_2 \cdot \mathrm{x}_i) + \cdots + c_k(\mathrm{x}_k \cdot \mathrm{x}_i) = \mathbf{0} \cdot \mathrm{x}_i = 0$$

Since, for $i \neq j$, $\mathrm{x}_j \cdot \mathrm{x}_i = 0$, we have

$$c_i(\mathbf{x}_i \cdot \mathbf{x}_i) = c_i \|\mathbf{x}_i\|^2 = 0 \quad (i = 1, 2, \ldots, k).$$

Since $\mathbf{x}_i \neq \mathbf{0}$ implies $\|\mathbf{x}_i\| > 0$ $(i = 1, 2, \ldots, k)$, we have

$$c_1 = c_2 = \cdots = c_k = 0.$$

Therefore, S is linearly independent. ■

Orthogonal Basis and Orthonormal Basis

Definition **[Orthonormal basis]**

Let S be a basis for $\mathbb{R}^n$. If S is orthogonal, then S is called an orthogonal basis. If S is orthonormal, then S is called an orthonormal basis.

Example 2 Sets in (1) and (3) of Example 1 are orthonormal bases of $\mathbb{R}^3$ and the set in (2) is an orthogonal basis of $\mathbb{R}^3$.

Theorem 7.7.2

Let $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be a basis for $\mathbb{R}^n$.

(1) If S is orthonormal, then each vector $\mathbf{x}$ in $\mathbb{R}^n$ can be expressed as

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n,$$

where $c_i = \mathbf{x} \cdot \mathbf{x}_i$ $(i = 1, 2, \cdots, n)$.

(2) If S is orthogonal, then $c_i = \dfrac{\mathbf{x} \cdot \mathbf{x}_i}{\|\mathbf{x}_i\|^2}$.

Proof We prove (1) only. Since S is a basis for $\mathbb{R}^n$, each vector $\mathbf{x} \in \mathbb{R}^n$ can be expressed as a linear combination of vectors in S as follows:

$$\mathrm{x} = c_1\mathrm{x}_1 + c_2\mathrm{x}_2 + \cdots + c_n\mathrm{x}_n, \quad (c_i \in \mathbb{R}).$$

For each i $(i = 1, 2, \ldots, n)$, we have

$$\begin{aligned}\mathrm{x}\cdot\ \mathrm{x}_i &= (c_1\mathrm{x}_1 + c_2\mathrm{x}_2 + \cdots + c_n\mathrm{x}_n)\cdot\ \mathrm{x}_i \\ &= c_1(\mathrm{x}_1\cdot\ \mathrm{x}_i) + c_2(\mathrm{x}_2\cdot\ \mathrm{x}_i) + \cdots + c_n(\mathrm{x}_n\cdot\ \mathrm{x}_i).\end{aligned}$$

Since S is orthonormal, $\mathrm{x}_i\cdot\ \mathrm{x}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$. Hence

$$c_i = \mathrm{x}\cdot\ \mathrm{x}_i \quad (i = 1, 2, \ldots, n).$$

■

Example 3

Write $\mathrm{y} = (2, -3, 5)$ as a linear combination of the vectors in

$$\left\{\mathrm{y}_1 = (0, 1, 0),\ \mathrm{y}_2 = \left(\frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}}\right),\ \mathrm{y}_3 = \left(\frac{1}{\sqrt{5}}, 0, -\frac{2}{\sqrt{5}}\right)\right\}$$

that is the orthonormal basis for $\mathbb{R}^3$ in (3)

Solution

Let $\mathrm{y} = c_1\mathrm{y}_1 + c_2\mathrm{y}_2 + c_3\mathrm{y}_3$. Then, by Theorem **7.7.2**, $c_i = \mathrm{y}\cdot\ \mathrm{y}_i$ $(i = 1, 2, 3)$.

Hence

$$c_1 = \mathrm{y}\cdot\ \mathrm{y}_1 = -3,\ c_2 = \mathrm{y}\cdot\ \mathrm{y}_2 = \frac{9}{\sqrt{5}},\ c_3 = \mathrm{y}\cdot\ \mathrm{y}_3 = -\frac{8}{\sqrt{5}}.$$

$$\therefore\ \mathrm{y} = c_1\mathrm{y}_1 + c_2\mathrm{y}_2 + c_3\mathrm{y}_3 = -3\mathrm{y}_1 + \frac{9}{\sqrt{5}}\mathrm{y}_2 - \frac{8}{\sqrt{5}}\mathrm{y}_3.$$

■

Theorem 7.7.3

(1) Suppose $S' = \{\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_n\}$ is an orthonormal basis for $\mathbb{R}^n$. Then, since $\|\mathrm{x}_i\| = 1$, the orthogonal projection $\mathrm{y} \in \mathbb{R}^n$ onto the subspace $W_k = <\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_k>$ in $\mathbb{R}^n$ is

$$\mathrm{y}_1 = \mathrm{proj}_{W_k}\mathrm{y} = (\mathrm{y}\cdot\ \mathrm{x}_1)\mathrm{x}_1 + (\mathrm{y}\cdot\ \mathrm{x}_2)\mathrm{x}_2 + \cdots + (\mathrm{y}\cdot\ \mathrm{x}_k)\mathrm{x}_k.$$

(2) If $S' = \{\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_n\}$ is an orthogonal basis, but not an orthonormal basis for $\mathbb{R}^n$, then $\mathrm{y}_1 = \mathrm{proj}_{W_k}\mathrm{y}$ can be written as

$$\mathbf{y}_1 = \mathrm{proj}_{W_k}\mathbf{y} = \frac{\mathbf{y}\cdot\ \mathbf{x}_1}{\|\mathbf{x}_1\|^2}\mathbf{x}_1 + \frac{\mathbf{y}\cdot\ \mathbf{x}_2}{\|\mathbf{x}_2\|^2}\mathbf{x}_2 + \cdots + \frac{\mathbf{y}\cdot\ \mathbf{x}_k}{\|\mathbf{x}_k\|^2}\mathbf{x}_k.$$

Example 4 Let W be a subspace of $\mathbb{R}^3$ spanned by the two vectors $\mathbf{x}_1 = (0,\ 1,\ 0)$, $\mathbf{x}_2 = \left(\frac{5}{13},\ 0,\ -\frac{12}{13}\right)$ in an orthonormal set $S = \{\mathbf{x}_1, \mathbf{x}_2\}$. Find the orthogonal projection of $\mathbf{y} = (2, 1, 1)$ onto W and the orthogonal component of $\mathbf{y}$ perpendicular to W.

Solution

$$\begin{aligned}\mathbf{y}_1 = \mathrm{proj}_W\mathbf{y} &= (\mathbf{y}\cdot\ \mathbf{x}_1)\mathbf{x}_1 + (\mathbf{y}\cdot\ \mathbf{x}_2)\mathbf{x}_2 \\ &= 1\,(0,\ 1,\ 0) + \left(-\frac{2}{13}\right)\left(\frac{5}{13},\ 0,\ -\frac{12}{13}\right) = \left(-\frac{10}{169},\ 1,\ \frac{24}{169}\right).\end{aligned}$$

The orthogonal component of $\mathbf{y}$ perpendicular to W is

$$\mathbf{y}_2 = \mathbf{y} - \mathrm{proj}_W\mathbf{y} = (2, 1, 1) - \left(-\frac{10}{169},\ 1,\ \frac{24}{169}\right) = \left(\frac{348}{169},\ 0,\ \frac{145}{169}\right).$$ ■

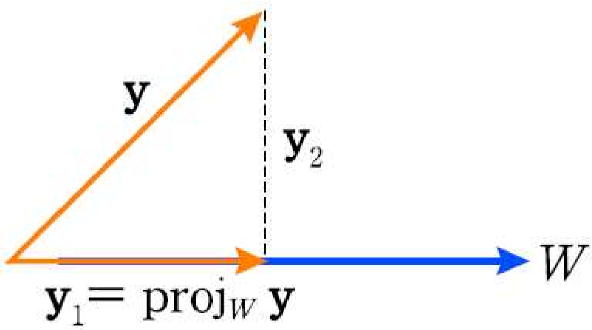
Gram-Schmidt orthonormal process

Theorem 7.7.4

Let $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be a basis for $\mathbb{R}^n$. Then we can obtain an orthonormal basis from S.

Proof [Gram-Schmidt Orthonormalization]

We first derive an orthogonal basis $T = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n\}$ for $\mathbb{R}^n$ from the basis S as follows:



[Step 1] Take $\mathbf{y}_1 = \mathbf{x}_1$.

[Step 2] Let W_1 be a subspace spanned by $\mathbf{y}_1$ and let

$$\mathbf{y}_2 = \mathbf{x}_2 - \text{proj}_{W_1}\mathbf{x}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{y}_1}{\|\mathbf{y}_1\|^2}\mathbf{y}_1.$$

[Step 3] Let W_2 be a subspace spanned by $\mathbf{y}_1$ and $\mathbf{y}_2$ and let

$$\mathbf{y}_3 = \mathbf{x}_3 - \text{proj}_{W_2}\mathbf{x}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{y}_1}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{y}_2}{\|\mathbf{y}_2\|^2}\mathbf{y}_2.$$

[Step 4] Repeat the same procedure to get

$$\mathbf{y}_k = \mathbf{x}_k - \text{proj}_{W_{k-1}}\mathbf{x}_k = \mathbf{x}_k - \frac{\mathbf{x}_k \cdot \mathbf{y}_1}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \frac{\mathbf{x}_k \cdot \mathbf{y}_2}{\|\mathbf{y}_2\|^2}\mathbf{y}_2 - \cdots - \frac{\mathbf{x}_k \cdot \mathbf{y}_{k-1}}{\|\mathbf{y}_{k-1}\|^2}\mathbf{y}_{k-1} \quad (k = 4, 5, \ldots, n),$$

where $W_k = < \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k >$.

It is clear that $T = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n\}$ is orthogonal. By taking

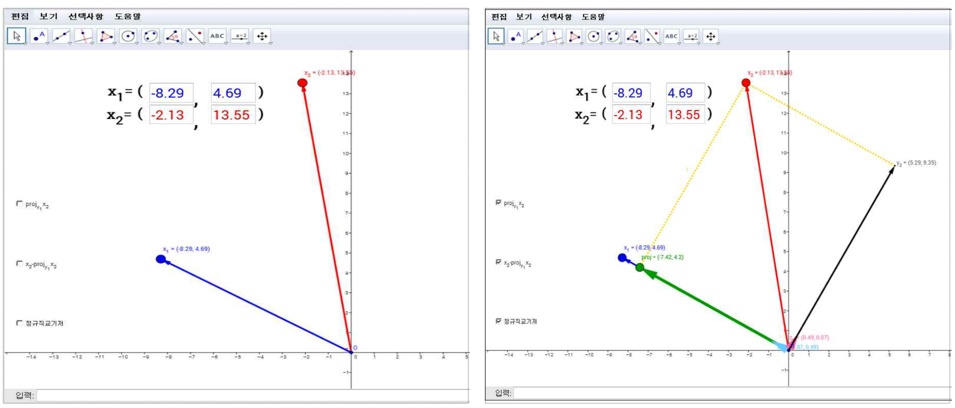
$$\mathbf{z}_k = \frac{\mathbf{y}_k}{\|\mathbf{y}_k\|} \quad (k = 1, 2, \ldots, n),$$

we get an orthonormal basis $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n\}$ for R^n. ■

The above process of producing and orthonormal basis from a given basis is called the Gram-Schmidt Orthogonalization process.

[Remark] **Simulation for Gram-Schmidt Orthonormalization**

- http://www.geogebratube.org/student/m58812



Example 5

Use the Gram-Schmidt Orthonormalization to find an orthonormal basis $Z = \{\mathbf{z}_1, \mathbf{z}_2\}$ for $\mathbb{R}^2$ from the two linearly independent vectors $\mathbf{x}_1 = (1, 1)$, $\mathbf{x}_2 = (1, 2)$.

Solution

We first find orthogonal vectors $\mathbf{y}_1$, $\mathbf{y}_2$ as follows:

[Step 1] $\mathbf{y}_1 = \mathbf{x}_1 = (1, 1)$

[Step 2] $\mathbf{y}_2 = \mathbf{x}_2 - \text{proj}_{W_1}\mathbf{x}_2 = \mathbf{x}_2 - \dfrac{\mathbf{x}_2 \cdot \mathbf{y}_1}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 = (1, 2) - \dfrac{3}{2}(1, 1) = \left(-\dfrac{1}{2}, \dfrac{1}{2}\right)$

$$\mathbf{z}_1 = \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \quad \mathbf{z}_2 = \frac{\mathbf{y}_2}{\|\mathbf{y}_2\|} = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \quad ■$$

Example 6

Let $\mathbf{x}_1 = (1, 1, 0)$, $\mathbf{x}_2 = (0, 1, 2)$, $\mathbf{x}_3 = (1, 2, 1)$. Use the Gram-Schmidt Orthonormalization to find an orthonormal basis $Z = \{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3\}$ for $\mathbb{R}^3$ using the basis $S = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ for $\mathbb{R}^3$.

Solution

We first find orthogonal vectors $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$:

[Step 1] Take $\mathbf{y}_1 = \mathbf{x}_1 = (1, 1, 0)$.

[Step 2]

$$\mathbf{y}_2 = \mathbf{x}_2 - \text{proj}_{W_1}\mathbf{x}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{y}_1}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 = (0, 1, 2) - \frac{1}{2}(1, 1, 0) = \left(-\frac{1}{2}, \frac{1}{2}, 2\right)$$

[Step 3] $\mathbf{y}_3 = \mathbf{x}_3 - \text{proj}_{W_2}\mathbf{x}_3 = \mathbf{x}_3 - \dfrac{\mathbf{x}_3 \cdot \mathbf{y}_1}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \dfrac{\mathbf{x}_3 \cdot \mathbf{y}_2}{\|\mathbf{y}_2\|^2}\mathbf{y}_2$

$$= (1, 2, 1) - \frac{3}{2}(1, 1, 0) - \frac{5}{9}\left(-\frac{1}{2}, \frac{1}{2}, 2\right) = \left(-\frac{2}{9}, \frac{2}{9}, -\frac{1}{9}\right)$$

By normalizing $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$, we get

$$\mathbf{z}_1 = \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$$

$$\mathbf{z}_2 = \frac{\mathbf{y}_2}{\|\mathbf{y}_2\|} = \left(-\frac{\sqrt{2}}{6}, \frac{\sqrt{2}}{6}, \frac{2\sqrt{2}}{3}\right)$$

$$\mathbf{z}_3 = \frac{\mathbf{y}_3}{\|\mathbf{y}_3\|} = \left(-\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right).$$

Therefore, $Z=\left\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(-\frac{\sqrt{2}}{6}, \frac{\sqrt{2}}{6}, \frac{2\sqrt{2}}{3}\right), \left(-\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right)\right\}$ □

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

① Computation for an orthogonal basis

```
x1=vector([1,1,0])
x2=vector([0,1,2])
x3=vector([1,2,1])
A=matrix([x1,x2,x3])          # generate a matrix with x1, x2, x3
[G,mu]=A.gram_schmidt()     # find an orthogonal basis. A==mu*G
print G
```

```
[   1     1     0]
[-1/2  1/2     2]
[-2/9  2/9 -1/9]
```

② Normalization

```
B=matrix([G.row(i) / G.row(i).norm() for i in range(0, 3)]); B
# The rows of matrix B are orthonormal
```

```
[     1/2*sqrt(2)        1/2*sqrt(2)                  0]
[-1/3*sqrt(1/2)   1/3*sqrt(1/2)   4/3*sqrt(1/2)]
[              -2/3                  2/3               -1/3]
```

Therefore, we get an orthonormal basis

$Z=\left\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(-\frac{\sqrt{2}}{6}, \frac{\sqrt{2}}{6}, \frac{2\sqrt{2}}{3}\right), \left(-\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right)\right\}$.

We can verify if Z is orthonormal as follows:

③ Checking for orthonormality

```
print B*B.transpose()       # Checking if B is an orthogonal matrix.
print
print B.transpose()*B
```

```
[1  0  0]            [1  0  0]
```

[0 1 0]　　　　[0 1 0]
[0 0 1]　　　　[0 0 1] ■

Example 7

Let $\mathbf{x}_1 = (1,\ 1,\ 2)$, $\mathbf{x}_2 = (0,\ 2,\ -4)$. Use the Gram-Schmidt Orthonormalization to find an orthonormal basis $Z = \{\mathbf{z}_1,\ \mathbf{z}_2\}$ for a subspace of $\mathbb{R}^3$ for which $S = \{\mathbf{x}_1,\ \mathbf{x}_2\}$ is a basis.

Solution

$\mathbf{y}_1 = \mathbf{x}_1 = (1,\ 1,\ 2)$

$\mathbf{y}_2 = \mathbf{x}_2 - \mathrm{proj}_{W_1}\mathbf{x}_2 = \mathbf{x}_2 - \dfrac{\mathbf{x}_2 \cdot \mathbf{y}_1}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 = (0,\ 2,\ -4) + \dfrac{6}{6}(1,\ 1,\ 2) = (1,\ 3,\ -2)$

$\therefore\ \mathbf{z}_1 = \dfrac{\mathbf{y}_1}{\|\mathbf{y}_1\|} = \left(\dfrac{1}{\sqrt{6}},\ \dfrac{1}{\sqrt{6}},\ \dfrac{2}{\sqrt{6}}\right),\ \ \mathbf{z}_2 = \dfrac{\mathbf{y}_2}{\|\mathbf{y}_2\|} = \left(\dfrac{1}{\sqrt{14}},\ \dfrac{3}{\sqrt{14}},\ -\dfrac{2}{\sqrt{14}}\right)$ ■

"Good, he did not have enough imagination to become a mathematician".

David Hilbert (1862-1943)
http://en.wikipedia.org/wiki/David_Hilbert
Hilbert is known as one of the founders of proof theory and mathematical logic,

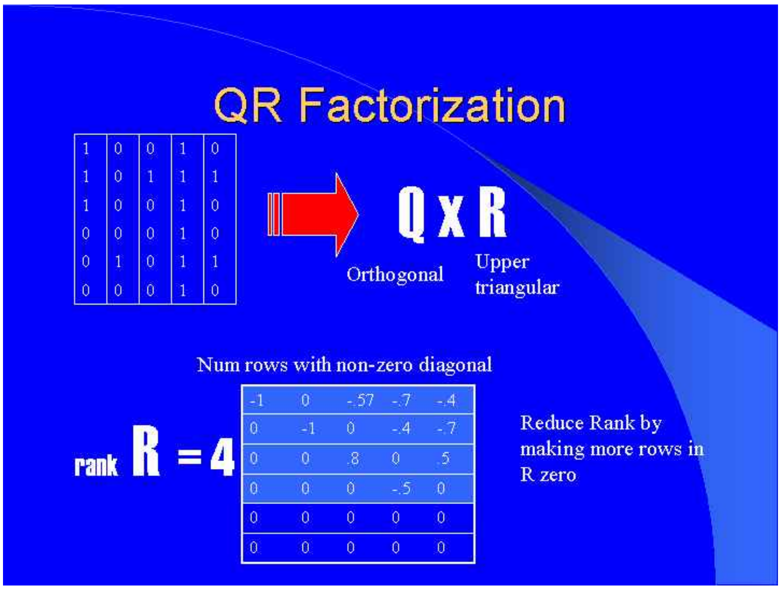
7.8 QR-Decomposition; Householder Transformations

- Lecture Movie : http://www.youtube.com/watch?v=crMXPi2lgGs
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-10-sec-7-8.html

If an $m \times k$ matrix A has k linearly independent columns, then the Gram-Schmidt Orthogonalization can be used to decompose the matrix A in the form of $A = QR$ where the columns of Q are the orthonormal vectors obtained by applying the Gram-Schmidt Orthognalization to the columns of A and R is an upper triangular matrix. The QR-decomposition is widely used to compute numerical solutions to linear systems, least-squares problems, and eigenvalue and eigenvector problems. In this section, we briefly introduce the QR-decomposition.

Details can be found in the following websites:

- http://www.math.ucla.edu/~yanovsky/Teaching/Math151B/handouts/GramSchmidt.pdf
- https://inst.eecs.berkeley.edu/~ee127a/book/login/l_mats_qr.html
- http://www.ugcs.caltech.edu/~chandran/cs20/qr.html



7.9 Coordinate vectors

- Lecture Movie : http://youtu.be/M4peLF7Xur0, http://youtu.be/tdd7gbtCCRg
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-10-sec-7-9.html

In a finite-dimensional vector space, a basis is closely related to a coordinate system. We have so far used the coordinate system associated to the standard basis of $\mathbb{R}^n$. In this section, we introduce coordinate systems based on non-standard bases. We also study the relationship between coordinate systems associated to different bases.

If $\alpha = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is an ordered basis for $\mathbb{R}^n$, then any vector $\mathbf{x}$ in $\mathbb{R}^n$ is uniquely expressed as a linear combination of the vectors in α as follows:

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n, \quad (c_1, c_2, \ldots, c_n \in \mathbb{R}) \tag{1}$$

Then $c_1, c_2, \ldots, c_n$ are called coordinates of the vector $\mathbf{x}$ relative to the basis α.

Definition **[Coordinate vectors]**

The scalars $c_1, c_2, \ldots, c_n$ in (1) are called the coordinates of $\mathbf{x}$ relative to the ordered basis α. Furthermore, the column vector in $\mathbb{R}^n$

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

is called the coordinate vector of $\mathbf{x}$ relative to the ordered basis α and denoted by $[\mathbf{x}]_\alpha$.

Example 1

The vector $\mathbf{x} = (2, -3, 5)$ in $\mathbb{R}^3$ can be expressed as follows relative to the standard basis $\alpha = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for $\mathbb{R}^3$:

$$\mathbf{x} = (2, -3, 5) = 2\mathbf{e}_1 + (-3)\mathbf{e}_2 + 5\mathbf{e}_3\,.$$

Therefore

$$[\mathbf{x}]_\alpha = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}.$$ ■

Example 2 Let $\mathbf{x}_1 = (1, 1, 0)$, $\mathbf{x}_2 = (1, 1, 1)$, $\mathbf{x}_3 = (0, 1, -1)$. For $\mathbf{x} = (1, 2, 3)$ find the coordinate vector $[\mathbf{x}]_\alpha$ relative to the basis $\alpha = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ for $\mathbb{R}^3$.

Solution

From $\mathbf{x} = (1, 2, 3) = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + c_3\mathbf{x}_3$, $(c_i \in \mathbb{R})$
$= c_1(1, 1, 0) + c_2(1, 1, 1) + c_3(0, 1, -1)$,

we get the linear system $\begin{cases} c_1 + c_2 = 1 \\ c_1 + c_2 + c_3 = 2 \\ c_2 - c_3 = 3 \end{cases}$.

By solving this linear system, we get $c_1 = -3$, $c_2 = 4$, $c_3 = 1$.

$$\therefore \; [\mathbf{x}]_\alpha = \begin{bmatrix} -3 \\ 4 \\ 1 \end{bmatrix}.$$ ■

- As described above, finding the coordinate vector relative to a basis is equivalent to solving a linear system.

Theorem 7.9.1

Let α be a basis for $\mathbb{R}^n$. For vectors $\mathbf{x}$, $\mathbf{y}$ in $\mathbb{R}^n$ and a scalar $c \in \mathbb{R}$, the following holds:

(1) $[\mathbf{x} + \mathbf{y}]_\alpha = [\mathbf{x}]_\alpha + [\mathbf{y}]_\alpha$.

(2) $[c\mathbf{x}]_\alpha = c[\mathbf{x}]_\alpha$.

- In general we have
$[c_1\mathbf{y}_1 + c_2\mathbf{y}_2 + \cdots + c_n\mathbf{y}_n]_\alpha = c_1[\mathbf{y}_1]_\alpha + c_2[\mathbf{y}_2]_\alpha + \cdots + c_n[\mathbf{y}_n]_\alpha$.

Change of Basis

Let $\alpha = \{\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_n\}$ and $\beta = \{\mathrm{y}_1, \mathrm{y}_2, \ldots, \mathrm{y}_n\}$ be two different ordered bases for $\mathbb{R}^n$. In the following, we consider a relationship between $[\mathrm{x}]_\alpha$ and $[\mathrm{x}]_\beta$.

- Letting $\mathrm{x} = c_1\mathrm{y}_1 + c_2\mathrm{y}_2 + \cdots + c_n\mathrm{y}_n$, $(c_i \in \mathbb{R})$, the coordinate vector of $\mathrm{x} \in \mathbb{R}^n$ relative to β is

$$[\mathrm{x}]_\beta = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix},$$

and the coordinate vector $[\mathrm{x}]_\alpha$ of $\mathrm{x} \in \mathbb{R}^n$ relative to α can be expressed as

$$[\mathrm{x}]_\alpha = [c_1\mathrm{y}_1 + c_2\mathrm{y}_2 + \cdots + c_n\mathrm{y}_n]_\alpha = c_1[\mathrm{y}_1]_\alpha + c_2[\mathrm{y}_2]_\alpha + \cdots + c_n[\mathrm{y}_n]_\alpha.$$

Let $[\mathrm{y}_j]_\alpha = \begin{bmatrix} p_{1j} \\ p_{2j} \\ \vdots \\ p_{nj} \end{bmatrix}$ be the coordiate vector of y_j relative to α and matrix P be

$$P = [[\mathrm{y}_1]_\alpha : [\mathrm{y}_2]_\alpha : \cdots : [\mathrm{y}_n]_\alpha] = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}.$$

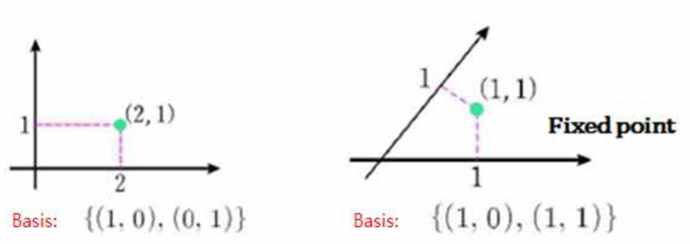
Then we have

$$\begin{aligned} [\mathrm{x}]_\alpha &= c_1 \begin{bmatrix} p_{11} \\ p_{21} \\ \vdots \\ p_{n1} \end{bmatrix} + c_2 \begin{bmatrix} p_{12} \\ p_{22} \\ \vdots \\ p_{n2} \end{bmatrix} + \cdots + c_n \begin{bmatrix} p_{1n} \\ p_{2n} \\ \vdots \\ p_{nn} \end{bmatrix} \\ &= \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = P[\mathrm{x}]_\beta, \end{aligned}$$

that is, $[\mathrm{x}]_\alpha = P[\mathrm{x}]_\beta$. (2)

- In the equation (2) matrix P transforms the coordinate vector $[\mathrm{x}]_\beta$ to another coordinate vector $[\mathrm{x}]_\alpha$. Hence the matrix $P = [[\mathrm{y}_1]_\alpha [\mathrm{y}_2]_\alpha \cdots [\mathrm{y}_n]_\alpha]$ is called a transition matrix from ordered basis β to ordered basis α and denoted by $P = [I]_\beta^\alpha$. Therefore, $[\mathrm{x}]_\alpha = P[\mathrm{x}]_\beta = [I]_\beta^\alpha [\mathrm{x}]_\beta$.

- This transformation is called change of basis. Note that the change of basis does not modify the nature of a vector, but it changes coordinate vectors. The following example illustrates this.



Example 3

Let $\alpha = \{\mathbf{e}_1, \mathbf{e}_2\}$ be the standard basis for $\mathbb{R}^2$ and $\mathbf{y}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{y}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. For the two different ordered bases α, $\beta = \{\mathbf{y}_1, \mathbf{y}_2\}$:

(1) Find the transition matrix $P = [I]_\beta^\alpha$ from basis β to basis α.

(2) Suppose $[\mathbf{x}]_\beta = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. Find the coordinate vector $[\mathbf{x}]_\alpha$.

(3) For $\mathbf{x} = \begin{bmatrix} 3 \\ 9 \end{bmatrix}$, show that equation (2) holds.

Solution

(1) Since $P = [I]_\beta^\alpha = [\,[\mathbf{y}_1]_\alpha \;\; [\mathbf{y}_2]_\alpha\,]$, we need to compute the coordinate vectors for $\mathbf{y}_1, \mathbf{y}_2$ relative to α. Since

$$\begin{cases} \mathbf{y}_1 = \mathbf{e}_1 + 2\mathbf{e}_2 \\ \mathbf{y}_2 = -\mathbf{e}_1 + \mathbf{e}_2 \end{cases},$$

$[\mathbf{y}_1]_\alpha = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $[\mathbf{y}_2]_\alpha = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Hence

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

(2) $[\mathbf{x}]_\alpha = P[\mathbf{x}]_\beta = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 7 \end{bmatrix}$

(3) Since $\mathbf{x} = \begin{bmatrix} 3 \\ 9 \end{bmatrix} = 3\mathbf{e}_1 + 9\mathbf{e}_2$ and also $\mathbf{x} = \begin{bmatrix} 3 \\ 9 \end{bmatrix} = 4\mathbf{y}_1 + 1\mathbf{y}_2$,

$[\mathbf{x}]_\alpha = \begin{bmatrix} 3 \\ 9 \end{bmatrix}$, $[\mathbf{x}]_\beta = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$.

It can be easily checked that $[\mathbf{x}]_\alpha = \begin{bmatrix} 3 \\ 9 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = P\,[\mathbf{x}]_\beta$ ■

Example **4**

For $\mathrm{x}_1 = (1, 2, 0)$, $\mathrm{x}_2 = (1, 1, 1)$, $\mathrm{x}_3 = (2, 0, 1)$ and
$\mathrm{y}_1 = (4, -1, 3)$, $\mathrm{y}_2 = (5, 5, 2)$, $\mathrm{y}_3 = (6, 3, 3)$,
let $\alpha = \{\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3\}$ and $\beta = \{\mathrm{y}_1, \mathrm{y}_2, \mathrm{y}_3\}$, both of which are bases for $\mathbb{R}^3$.
Find $P = [I]_\beta^\alpha$.

Solution

Since $P = [[\mathrm{y}_1]_\alpha\ [\mathrm{y}_2]_\alpha\ [\mathrm{y}_3]_\alpha]$, we first find the coordinate vectors for y_1, y_2, y_3 relative to α. Letting

$$a_1\mathrm{x}_1 + a_2\mathrm{x}_2 + a_3\mathrm{x}_3 = \mathrm{y}_1,\ (a_i \in \mathbb{R})$$
$$b_1\mathrm{x}_1 + b_2\mathrm{x}_2 + b_3\mathrm{x}_3 = \mathrm{y}_2,\ (b_i \in \mathbb{R})$$
$$c_1\mathrm{x}_1 + c_2\mathrm{x}_2 + c_3\mathrm{x}_3 = \mathrm{y}_3,\ (c_i \in \mathbb{R}),$$

we get the following three linear systems:

$$\begin{aligned} a_1 + a_2 + 2a_3 &= 4 \\ 2a_1 + a_2 \quad &= -1 \\ a_2 + a_3 &= 3 \end{aligned} \qquad \begin{aligned} b_1 + b_2 + 2b_3 &= 5 \\ 2b_1 + b_2 \quad &= 5 \\ b_2 + b_3 &= 2 \end{aligned} \qquad \begin{aligned} c_1 + c_2 + 2c_3 &= 6 \\ 2c_1 + c_2 \quad &= 3 \\ c_2 + c_3 &= 3 \end{aligned}$$

Note that all of the above linear systems have $A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ as their coefficient matrix. Hence we can solve the linear systems simultaneously using the RREF of the coefficient matrix. That is, by converting the augmented matrix $[A \vdots\ \mathrm{y}_1 \vdots\ \mathrm{y}_2 \vdots\ \mathrm{y}_3]$ in its RREF, we can find the values of a_i, b_i, $c_i (i = 1, 2, 3)$ at the same time:

$$A = \left[\begin{array}{ccc:c:c:c} 1 & 1 & 2 & 4 & 5 & 6 \\ 2 & 1 & 0 & -1 & 5 & 3 \\ 0 & 1 & 1 & 3 & 2 & 3 \end{array}\right]$$

has the RREF

$$B = \left[\begin{array}{ccc:c:c:c} 1 & 0 & 0 & -1 & 2 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 & 1 & 2 \end{array}\right].$$

Therefore, the transition matrix from β to α is

$$P = [I]_\beta^\alpha = \begin{bmatrix} -1 & 2 & 1 \\ 1 & 1 & 1 \\ 2 & 1 & 2 \end{bmatrix}. \quad \square$$

- http://matrix.skku.ac.kr/RPG_English/7-MA-transition-matrix.html

Sage http://sage.skku.edu

```
x1=vector([1,2,0]);x2=vector([1,1,1]);x3=vector([2,0,1])
A=column_matrix([x1, x2, x3])
y1=vector([4, -1, 3]);y2=vector([5, 5, 2]);y3=vector([6, 3, 3])
B=column_matrix([y1, y2, y3]) # Creating the matrix with columns y1, y2,
                                # y3
aug=A.augment(B, subdivide=True)
aug.rref()
```

```
[ 1  0  0|-1  2  1]
[ 0  1  0| 1  1  1]
[ 0  0  1| 2  1  2]
```

■

Theorem 7.9.2

Suppose α and β are two different ordered bases for $\mathbb{R}^n$ and P be the transition matrix from β to α. Then P is invertible and its inverse $P^{-1}$is the transisiton matrix from α to β, that is, $P^{-1}=[I]_{\alpha}^{\beta}$.

Example 5

For the two bases α, β for $\mathbb{R}^3$ in Example 4, compute the following:

(1) The transition matrix $Q=[I]_{\alpha}^{\beta}$ from basis α to basis β.

(2) The coordinate vector $[\mathbf{x}]_{\beta}$ relative to basis β for given $[\mathbf{x}]_{\alpha}=\begin{bmatrix}1\\5\\2\end{bmatrix}$.

Solution

(1) Since the transition matrix from β to α is $P=\begin{bmatrix}-1&2&1\\1&1&1\\2&1&2\end{bmatrix}$, by Theorem 7.9.2, we have

$$Q=[I]_\alpha^\beta=P^{-1}=\begin{bmatrix}-\frac{1}{2} & \frac{3}{2} & -\frac{1}{2}\\ 0 & 2 & -1\\ \frac{1}{2} & -\frac{5}{2} & \frac{3}{2}\end{bmatrix}.$$

(2) $[\mathbf{x}]_\beta=Q[\mathbf{x}]_\alpha=\begin{bmatrix}-\frac{1}{2} & \frac{3}{2} & -\frac{1}{2}\\ 0 & 2 & -1\\ \frac{1}{2} & -\frac{5}{2} & \frac{3}{2}\end{bmatrix}\begin{bmatrix}1\\5\\2\end{bmatrix}=\begin{bmatrix}6\\8\\-9\end{bmatrix}.$

Sage http://sage.skku.edu

```
x1=vector([1,2,0]);x2=vector([1,1,1]);x3=vector([2,0,1])
x0=vector([1,5,2])
A=column_matrix([x1, x2, x3])
y1=vector([4, -1, 3]);y2=vector([5, 5, 2]);y3=vector([6, 3, 3])
B=column_matrix([y1, y2, y3])
aug=B.augment(A, subdivide=True)
Q=aug.rref()
print Q
```

```
[   1    0    0|-1/2  3/2 -1/2]
[   0    1    0|   0    2   -1]
[   0    0    1| 1/2 -5/2  3/2]
```

■

[Bookmarks] http://blog.daum.net/with-learn/5432044

Chapter 7 Exercises

- http://matrix.skku.ac.kr/LA-Lab/index.htm
- http://matrix.skku.ac.kr/knou-knowls/cla-sage-reference.htm

Problem 1 Use determinant to check if the following vectors are linearly independent:
$\mathrm{v}_1 = (1,\ 1,\ -3),\ \mathrm{v}_2 = (0,\ 2,\ 1),\ \mathrm{v}_3 = (0,\ -1,\ 0)$

Problem 2 Determine if the given set S is a basis for $\mathbb{R}^3$.

(1) $S_1 = \{(2,0,1),\ (6,3,5),\ (0,-5,0)\}$

(2) $S_2 = \{(0,\ 2,\ 3),\ (2,\ 4,\ 1),\ (1,\ 3,\ 2)\}$
(Hint: http://math3.skku.ac.kr/spla/CLA-7.1-Exercise-2)

Problem 3 Find two different bases for the subspace of $\mathbb{R}^3$ described by the equation $x+2y+3z=0$.

Problem 4 Given a homogeneous linear system, find a basis and the dimension of its corresponding solution space.

(1)
$$\begin{aligned} 4x_1 - 2x_2 + x_3 + 2x_4 &= 0 \\ x_1 + x_2 + x_3 - x_4 &= 0 \end{aligned}$$
(Hint: http://math1.skku.ac.kr/home/matrix1/261/)

(2)
$$\begin{aligned} x_1 + 3x_2 - 2x_3 + 2x_5 &= 0 \\ 2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 &= 0 \\ 5x_3 + 10x_4 + 15x_6 &= 0 \\ 2x_1 + 6x_2 + 8x_4 + 4x_5 + 18x_6 &= 0 \end{aligned}$$
(Hint: http://math1.skku.ac.kr/home/pub/548/)

Problem 5 Given matrix A, find a basis for its null space and nullity(A).

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

Solution Sage: Find RREF of $[A\,\vdots\ \ 0]$

```
A=matrix(ZZ,6,5,[2,2,-1,0,1,-1,-1,2,-3,1,1,1,-2,0,-1,0,0,1,1,1,0,0,0,1,1,0,0,1,1,0])
A.echelon_form()
A.right_kernel()
```

$\begin{bmatrix} 1\,1\,0\,0\,0:0 \\ 0\,0\,1\,0\,0:0 \\ 0\,0\,0\,1\,0:0 \\ 0\,0\,0\,0\,1:0 \\ 0\,0\,0\,0\,0:0 \\ 0\,0\,0\,0\,0:0 \end{bmatrix}$. $\mathbf{x}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}=\begin{bmatrix} -x_2 \\ x_2 \\ 0 \\ 0 \\ 0 \end{bmatrix}=s\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $(s\in R)$ forms a set of solutions. $S=\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}\right\}$ and nullity(A) = 1. ■

Problem 6 For the following matrix A, find a basis for its column space $\mathrm{Col}(A)$ and compute the column rank $c(A)$.

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

Problem 7 For given matrix A compute its rank and nullity. Verify if the rank and nullity of A satisfy the Rank-Nullity Theorem.

(1) $A=\begin{bmatrix} 2 & 5 & 7 & 9 & 10 & 11 \\ 2 & 3 & 1 & 2 & 4 & 8 \\ 8 & 6 & 2 & 1 & 2 & 3 \end{bmatrix}$

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

Solution ① RREF(A)

```
A=matrix(QQ, 3, 6, [2, 5, 7, 9, 10, 11, 2, 3, 1, 2, 4, 8, 8, 6, 2, 1, 2, 3])
print A.echelon_form()
```

```
[    1     0     0  -3/4  -3/2  -13/4]
[    0     1     0   7/8   9/4   21/4]
[    0     0     1   7/8   1/4   -5/4]
```

=> rank(A)=3 and nulllity (A)= 6-3 =3.

② Sage:

```
A=matrix(QQ, 3, 6, [2, 5, 7, 9, 10, 11, 2, 3, 1, 2, 4, 8, 8, 6, 2, 1, 2, 3])
print A.rank()
print A.right_nullity()
```

∴ rank(A)=3 and nullity(A)=3. ■

Problem 8 Check if $\text{rank}(A) = \text{rank}(A^T)$.

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

Problem 9 Using the table below compute $\text{Row}(A), \text{Null}(A)$, $\text{Col}(A), \text{Null}(A^T)$ for matrix A:

	(a)	(b)	(c)	(d)	(e)
size of A	3×3	3×3	3×3	5×9	9×5
$\text{rank}(A)$	3	2	1	3	2

Problem 10 For $A \in M_{m\times n}$, if $\text{rank}(A) = m$, we say that A has full row ran, and if $\text{rank}(A) = n$, A is said to have full column rank. Determine if A has full row rank and/or full column rank:

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

(Hint: http://math1.skku.ac.kr/home/pub/565/)

Problem 11 For $\mathbf{a} = (1, 2, -1, 1)$, find a basis and the dimension of the hyperplane $\mathbf{a}^\perp = \{\mathbf{x} \mid \mathbf{a}\cdot \mathbf{x} = 0\}$.

Solution For any $\mathbf{x}=(x,\ y,\ z,\ w)$ in $\mathbf{a}^{\perp}$, $(1,2,-1,1)\cdot\ (x,y,z,w)=0$.

=> $x+2y-z+w=0$ => $\dim\ \mathbf{a}^{\perp}\ =4-1=3$ (since $\mathbf{a}\in R^4$).

A set {(1, 0, 0, -1), (0, 1, 0 ,-2), (0, 0, 1, 1)} forms a basis for hyperplane $\mathbf{a}^{\perp}$. ■

Problem 12 For $\mathbf{x}=(1,\ 2,\ 1)$ and $\mathbf{a}=(2,\ 1,\ -1)$, find the standard matrix for $T(\mathbf{x})=\mathrm{proj}_{<\mathbf{a}>}\mathbf{x}$.

Problem 13 For $\mathbf{x}=(1,\ 2,\ 1)$ and $\mathbf{a}=(2,\ 1,\ -1)$, using $\mathrm{proj}_{<\mathbf{a}>}\mathbf{x}$, find $\mathbf{x}_1$ and $\mathbf{x}_2$ such that $\mathbf{x}_1\in<\mathbf{a}>$, $\mathbf{x}_2\in<\mathbf{a}>^{\perp}$ and $\mathbf{x}=\mathbf{x}_1+\mathbf{x}_2$.

Solution $T(\mathbf{x})=\mathrm{proj}_{<\mathbf{a}>}\mathbf{x}=P\mathbf{x}$ and its standard matrix is

$$P=\frac{1}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T=\frac{1}{6}\begin{bmatrix}2\\1\\-1\end{bmatrix}[2\ 1\ -1]=\frac{1}{6}\begin{bmatrix}4&2&-2\\2&1&-1\\-2&-1&1\end{bmatrix}.$$

Since $\mathbf{x}_1\in<\mathbf{a}>$ and $\mathbf{x}_1=T(\mathbf{x})$, $\mathbf{x}_1=P\mathbf{x}=\frac{1}{6}\begin{bmatrix}4&2&-2\\2&1&-1\\-2&-1&1\end{bmatrix}\begin{bmatrix}1\\2\\1\end{bmatrix}=\frac{1}{6}\begin{bmatrix}6\\3\\-3\end{bmatrix}=\begin{bmatrix}1\\\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}$.

$$\therefore\ \mathbf{x}_2=\mathbf{x}-\mathbf{x}_1=\begin{bmatrix}1\\2\\1\end{bmatrix}-\begin{bmatrix}1\\\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}=\begin{bmatrix}0\\\frac{3}{2}\\\frac{3}{2}\end{bmatrix}.$$ ■

Problem 14 For given $\mathbf{x}=(1,\ 0,\ 2)$, express $\mathbf{x}$ as $\mathbf{x}=\mathbf{x}_1+\mathbf{x}_2$ for which $\mathbf{x}_1$ is in the direction of $\mathbf{a}=(2,\ 3,\ 1)\mathbf{a}$ and $\mathbf{x}_2$ is perpendicular to $\mathbf{a}$.

Problem 15 For the following A and $\mathbf{b}$, find the least squares solution to $A\mathbf{x}=\mathbf{b}$:

(1) $A=\begin{bmatrix}1&3&5&2&-2\\-2&1&-2&1&0\\-1&-3&0&1&0\\-3&0&1&0&0\\0&1&2&-2&1\end{bmatrix}$, $\mathbf{b}=\begin{bmatrix}1\\0\\1\\0\\0\end{bmatrix}$.

(2) $A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 4 & 3 & 2 & 1 \\ -1 & 2 & 3 & 4 & 5 & 4 & 3 & 2 & -1 \\ 1 & -2 & 3 & 4 & 5 & 4 & 3 & -2 & 1 \\ 1 & 2 & -3 & 4 & 5 & 4 & -3 & 2 & 1 \\ 1 & 2 & 3 & -4 & 5 & -4 & 3 & 2 & 1 \\ 1 & 2 & 3 & 4 & -5 & 4 & 3 & 2 & 1 \end{bmatrix}$, $\mathrm{b} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \\ 0 \\ 2 \end{bmatrix}$.

Problem 16 Find the least squares curve $y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$ passing through the five points $(1, 5)$, $(2, 1)$, $(3, -3)$, $(4, 1)$, $(5, 2)$.

Problem 17 Determine the values of a, b, c which make the set $\{(1, 1, 1), (3, 2, -5), (a, b, c)\}$ orthogonal.

Problem 18 Find the orthonormal set relative to the following orthogonal vectors: $\mathrm{v}_1 = (1\ 2, 1)$, $\mathrm{v}_2 = (1, 0, 1)$, $\mathrm{v}_3 = (3, 1, 0)$.

Solution Sage:

```
x1=vector([1,2,1])
x2=vector([1,0,1])
x3=vector([3,1,0])
A=matrix([x1,x2,x3])
[G,mu]=A.gram_schmidt()
B=matrix([G.row(i) / G.row(i).norm() for i in range(0, 3)]); B
```

```
[1/6*sqrt(6) 1/3*sqrt(6) 1/6*sqrt(6)]
[  sqrt(1/3)  -sqrt(1/3)   sqrt(1/3)]
[  sqrt(1/2)           0  -sqrt(1/2)] .
```
■

Problem 19 Show that each of the following sets of vectors $\mathbb{R}^4$ is linearly independent, and find its corresponding orthonormal set:

(1) $\mathrm{v}_1 = (0, 0, 1, 0)$, $\mathrm{v}_2 = (1, 0, 1, 1)$, $\mathrm{v}_3 = (1, 1, 2, 1)$.

(2) $\mathrm{v}_1 = (1, 1, 1, 1)$, $\mathrm{v}_2 = (-1, 4, 4, -1)$, $\mathrm{v}_3 = (4, -2, 2, 0)$.

Problem 20 For given plane $\pi: \{(x,y,z) | x - y + 2z = 0\}$ and vector $\mathbf{v} = (2,4,-3)$, find the following (Note that the inner product is defined to be $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v}$.):

(1) A basis for the 2-dimensional vector space represented by the plane and its corresponding orthonormal basis

(2) $\mathrm{proj}_{\pi}\mathbf{v}$

Problem 21 For the ordered basis $\beta = \{(0,1,1,1), (1,0,-1,1), (1,2,0,2), (3,-2,2,0)\}$ for $\mathbb{R}^4$:

(1) For $\mathbf{x} = (7,-7,5,4)$, find its coordinate vector $[\mathbf{x}]_\beta$ relative to β.

(2) For $\mathbf{y} = (1,-4,4,3)$, find its coordinate vector $[\mathbf{y}]_\beta$ relative to β.

(3) Find the coordinate vector $[2\mathbf{x}+\mathbf{y}]_\beta$ of $2\mathbf{x}+\mathbf{y}$ relative to β.

(4) For the above $\mathbf{x}$ and $\mathbf{y}$, find $[3\mathbf{x}]_\beta$ and $[-5\mathbf{y}]_\beta$.

Solution (1) $\begin{bmatrix} \frac{17}{2} \\ \frac{29}{4} \\ \frac{47}{8} \\ \frac{15}{8} \end{bmatrix}$ (2) $\begin{bmatrix} \frac{19}{2} \\ \frac{25}{4} \\ \frac{51}{8} \\ \frac{3}{8} \end{bmatrix}$.

Sage :

```
x1=vector([0,1,1,1]) ; x2=vector([1,0,-1,1]); x3=vector([1,2,0,2]); x4=vector([3,-2,2,0])
P=column_matrix([x1,x2,x3,x4])
A1=matrix(4,1,[7, -7, 5, 4]); A2=matrix(4,1,[1, -4, 4, 3])
P1=P.inverse()
print P1*A1; print; print P1*A2
```
■

Problem 22 For $\mathbf{u}_1 = (1, 2)$, $\mathbf{u}_2 = (2, 3)$, $\mathbf{v}_1 = (1, 3)$, $\mathbf{v}_2 = (1, 4)$, let $\alpha = \{\mathbf{u}_1, \mathbf{u}_2\}$, $\beta = \{\mathbf{v}_1, \mathbf{v}_2\}$ which are bases for $\mathbb{R}^2$.

(1) Find the transition matrix $[I]_\beta^\alpha$.

(2) Find the transition matrix $[I]_\alpha^\beta$.

(3) Suppose $[\mathbf{w}]_\alpha = (1,\ 1)$. Find $[\mathbf{w}]_\beta$ using the transition matrix $[I]_\alpha^\beta$.

(4) Suppose $[\mathbf{w}]_\beta = (3,\ 2)$. Find $[\mathbf{w}]_\alpha$ using the transition matrix $[I]_\beta^\alpha$.

Problem P1 If the size of matrix A is $m \times n$, what is the value of $\operatorname{rank}(A^T) + \operatorname{nullity}(A^T)$?

Problem P2 **(Select one)** If one replaces a matrix with its transpose, then

A. The image may change, but the kernel, rank, and nullity do not change.
B. The image, kernel, rank, and nullity may all change.
C. The image, rank, and kernel may change, but the nullity does not change.
D. The image, kernel, rank, and nullity all do not change.
E. The image, kernel, and nullity may change, but the rank does not change.
F. The kernel may change, but the image, rank, and nullity do not change.
G. The image and kernel may change, but the rank and nullity do not change.

Problem P3 **(Select one)** Let $T: \mathbb{R}^3 \to \mathbb{R}^5$ be a linear transformation. Then

A. T is invertible if and only if the rank is five.
B. T is one-to-one if and only if the rank is three; T is never onto.
C. T is onto if and only if the rank is two; T is never one-to-one.
D. T is one-to-one if and only if the rank is two; T is never onto.
E. T is onto if and only if the rank is three; T is never one-to-one.
F. T is onto if and only if the rank is five; T is never one-to-one.
G. T is one-to-one if and only if the rank is five; T is never onto.

Problem P4 **(Select one)** If a linear transformation $T: \mathbb{R}^3 \to \mathbb{R}^5$ is onto, then

A. The rank is three and the nullity is zero.
B. The rank and nullity can be any pair of non-negative numbers that add up to three.
C. The rank is three and the nullity is two.
D. The rank is two and the nullity is three.
E. The situation is impossible.
F. The rank and nullity can be any pair of non-negative numbers that add up to five.
G. The rank is five and the nullity is two.

Problem P5 **(Select one)** If a linear transformation $T: \mathbb{R}^3 \to \mathbb{R}^5$ is one-to-one, then

A. The rank is five and the nullity is two.
B. The situation is impossible.
C. The rank and nullity can be any pair of non-negative numbers that add up to five.
D. The rank is two and the nullity is three.
E. The rank is three and the nullity is zero.
F. The rank is three and the nullity is two.
G. The rank and nullity can be any pair of non-negative numbers that add up to three.

Problem P6

(1) If the homogeneous linear system $A\mathbf{x} = \mathbf{0}$ has m linear equations and n unknowns, what is the maximum possible value for the dimension of the solution space?

(2) What is the dimension of a hyperplane in $\mathbb{R}^6$ perpendicular to a vector $\mathbf{a}$ in $\mathbb{R}^6$?

(3) List all of the possible dimensions of subspaces of $\mathbb{R}^5$?

(4) What is the dimension of the subspace of $\mathbb{R}^4$ spanned by the three vectors $\mathbf{v}_1 = (1,0,1,0)$, $\mathbf{v}_2 = (1,1,0,0)$, $\mathbf{v}_3 = (1,1,1,0)$?

Problem P7 Suppose $S = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is a basis for $\mathbb{R}^n$. If A is an invertible matrix of order n, show that the set $\{A\mathbf{x}_1, \ldots, A\mathbf{x}_n\}$ is also a basis for $\mathbb{R}^n$.

Problem P8 Determine if the following matrix A and A^TA have the same null space and row space:

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

Problem P9 What happens if the Gram-Schmidt Orthonormalization Procedure is applied to linearly dependent vectors?

Problem P10 Suppose the columns of A are orthonormal. What is a relationship between the column spaces of AA^T and A?

Problem PI1 Show that the set $S=\{(2,1,0),(0,0,2),(4,1,5)\}$ spans $\mathbb{R}^3$.

Problem PI2 What are the possible ranks of A according to the varying values of t:

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

[2014 ICM Seoul, Korea] http://www.icm2014.org/

Chapter 8 Diagonalization

In Chapter 6, we have studied how to represent a linear transformation from $\mathbb{R}^n$ into $\mathbb{R}^m$ as a matrix using its corresponding standard matrix. We were able to compute the standard matrix of the linear transformation based on the fact that every vector in $\mathbb{R}^n$ or $\mathbb{R}^m$ can be expressed as a linear combination of the standard basis vectors.

In this chapter, we study how to represent a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ with respect to arbitrary ordered bases for $\mathbb{R}^n$ and $\mathbb{R}^m$. In addition, we study relationship between different matrix representations of a linear transformation from $\mathbb{R}^n$ to itself using transition matrices. We also study matrix diagonalization.

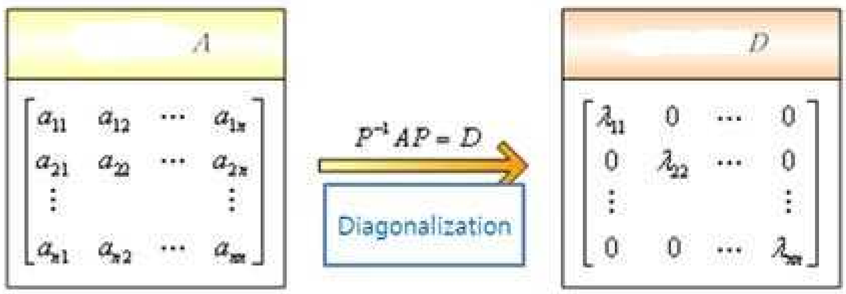
Further we study spectral properties of symmetric matrices and show that every symmetric matrix is orthogonally diagonalizable.

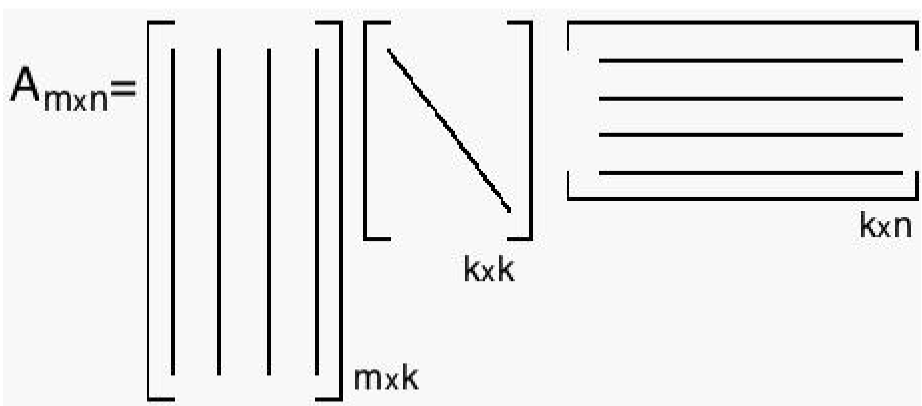
* A quadratic form is a quadratic equation which we come across in mathematics, physics, economics, statistics, and image processing, etc. Symmetric matrices play a significant role in the study of quadratic forms. In particular, we will learn how orthogonal diagonalization of symmetric matrices is used in the study of quadratic forms.

We introduce one of the most important concept in matrix theory called the singular value decomposition (SVD) which find many applications in science and engineering.

We will generalize matrix diagonalization of $m \times n$ matrices and study least squares solutions and a pseudoinverse.

We introduce complex matrices having complex eigenvalues and eigenvectors. We also introduce Hermitian matrices and unitary matrices that are complex counterparts corresponding to symmetric matrices and orthogonal matrices, respectively. Lastly, we study diagonalization of complex matrices.



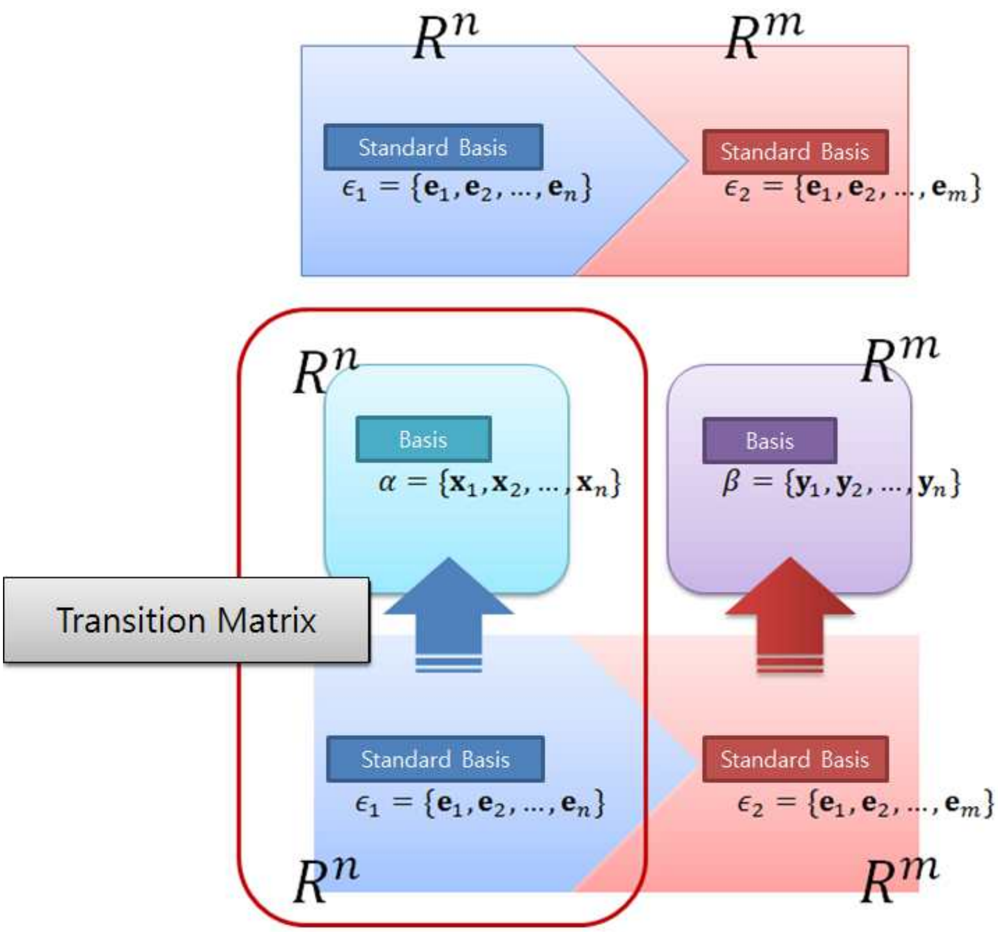


8.1 Matrix Representation

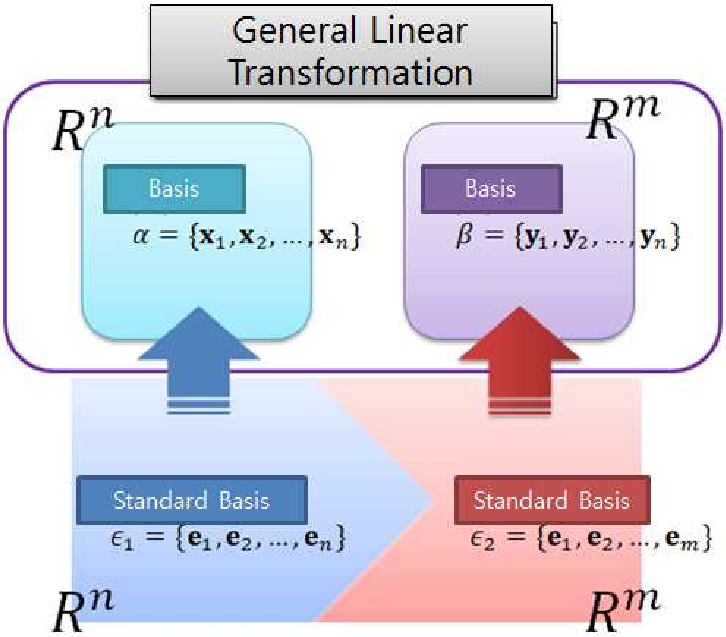
- Lecture Movie : http://youtu.be/jfMcPoso6g4
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-11-sec-8-1.html

In Chapter 6, we have studied how to represent a linear transfromation from $\mathbb{R}^n$ into $\mathbb{R}^m$ as a matrix using the standard bases for $\mathbb{R}^n$ and $\mathbb{R}^m$. In this section, we find a matrix representation of a linear transformation from $\mathbb{R}^n$ into $\mathbb{R}^m$ with respect to arbitrary ordered bases for $\mathbb{R}^n$ and $\mathbb{R}^m$.

Matrix Representation Relative to the Standard Bases



Matrix Representation In Some Ordered Bases for $\mathbb{R}^n$ and $\mathbb{R}^m$



Theorem 8.1.1

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let

$$\alpha = \{\mathrm{x}_1, \ldots, \mathrm{x}_n\}, \quad \beta = \{\mathrm{y}_1, \ldots, \mathrm{y}_m\}$$

be ordered bases for R^n and R^m, respectively. Let $\mathrm{y} = T(\mathrm{x})$. Then

$$[\mathrm{y}]_\beta = A' \; [\mathrm{x}]_\alpha = [T]_\alpha^\beta \; [\mathrm{x}]_\alpha ,$$

where the matrix representation $A' = [T]_\alpha^\beta$ of T in the ordered bases α and β is

$$A' = [[T(\mathrm{x}_1)]_\beta : [T(\mathrm{x}_2)]_\beta : \cdots : [T(\mathrm{x}_n)]_\beta].$$

Note that the matrix $[T]_\alpha^\beta$ is called the matrix associated with the linear transformation T with respect to the bases α and β.

Proof Recall that any vector $\mathrm{x} \in \mathbb{R}^n$ can be uniquely represented as a linear combination of vectors in $\alpha = \{\mathrm{x}_1, \ldots, \mathrm{x}_n\}$, say

$$\mathrm{x} = c_1\mathrm{x}_1 + c_2\mathrm{x}_2 + \cdots + c_n\mathrm{x}_n.$$

Then the coordinate vector for $\mathbf{x}$ relative to the basis α is

$$[\mathrm{x}]_\alpha = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

By the linearity of T, we have $\mathbf{y} = T(\mathrm{x}) = c_1 T(\mathrm{x}_1) + c_2\, T(\mathrm{x}_2) + \cdots + c_n\, T(\mathrm{x}_n)$. Since $\mathbf{y}$ is a vector in $\mathbb{R}^m$, the coordinate vector of $\mathbf{y}$ relative to β satisfies

$$\begin{aligned}[\mathrm{y}]_\beta &= c_1 [T(\mathrm{x}_1)]_\beta + c_2\, [T(\mathrm{x}_2)]_\beta + \cdots + c_n\, [T(\mathrm{x}_n)]_\beta \\ &= [[T(\mathrm{x}_1)]_\beta : [T(\mathrm{x}_2)]_\beta : \ \cdots : \ [T(\mathrm{x}_n)]_\beta] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = A'\ [\mathrm{x}]_\alpha \end{aligned}$$ ■

Thus the matrix $[T]_\alpha^\beta$ is the matrix whose ith column is the coordinate vector $[T(\mathrm{x}_i)]_\beta$ of $T(\mathrm{x}_i)$ with respect to the basis β.

[Remarks]

(1) By **Theorem 8.1.1** we can compute $[T(\mathrm{x})]_\beta$ by a matrix-vector multiplication, that is,

$$[\mathrm{y}]_\beta = [T(\mathrm{x})]_\beta = [T]_\alpha^\beta [\mathrm{x}]_\alpha = A' [\mathrm{x}]_\alpha.$$

(2) The matrix $[T]_\alpha^\beta = A'$ varies according the ordered bases α, β. For example, if we change the order of the vectors in the ordered base α, then the columns of A' change as well.

(3) The matrices $A = [T] = [T]_{\epsilon_1}^{\epsilon_2}$ and $A' = [T]_\alpha^\beta$ are distinct, but they have the following relationship:

$$A' = [T]_\alpha^\beta = [I]_{\epsilon_2}^\beta\, [T]_{\epsilon_1}^{\epsilon_2} [I]_\alpha^{\epsilon_1} \cong [T]_{\epsilon_1}^{\epsilon_2} = A$$

(4) If $R^n = R^m$ and $\alpha = \beta$, then $A' = [T]_\alpha^\alpha$ is denoted by $A = [T]_\alpha$ and is called the matrix representation of T relative to the ordered basis α.

Example 1

Define a linear transformation $T: \mathbb{R}^3 \to \mathbb{R}^2$ via $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 2x - z \\ y - z \end{bmatrix}$ and let

$$\alpha = \left\{ \mathrm{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},\ \mathrm{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},\ \mathrm{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\},\ \beta = \left\{ \mathrm{y}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},\ \mathrm{y}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$$

be ordered bases for $\mathbb{R}^3$ and $\mathbb{R}^2$, respectively. Compute $A' = [T]_\alpha^\beta$.

Solution

Since $A' = [T]_\alpha^\beta = [[T(\mathrm{x}_1)]_\beta\ [T(\mathrm{x}_2)]_\beta\ [T(\mathrm{x}_3)_\beta]]_{2\times 3}$, we first compute $T(\mathrm{x}_1),\ T(\mathrm{x}_2),\ T(\mathrm{x}_3)$:

$$T(\mathrm{x}_1) = \begin{bmatrix} 1 \\ -1 \end{bmatrix},\ T(\mathrm{x}_2) = \begin{bmatrix} -1 \\ 0 \end{bmatrix},\ T(\mathrm{x}_3) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

We now find the coordinate vectors of the above vectors relative to the ordered basis β. Since

$$\begin{aligned} T(\mathrm{x}_1) &= \begin{bmatrix} 1 \\ -1 \end{bmatrix} = a_1\mathrm{y}_1 + a_2\mathrm{y}_2 = a_1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + a_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \\ T(\mathrm{x}_2) &= \begin{bmatrix} -1 \\ 0 \end{bmatrix} = b_1\mathrm{y}_1 + b_2\mathrm{y}_2 = b_1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + b_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \\ T(\mathrm{x}_3) &= \begin{bmatrix} 2 \\ 1 \end{bmatrix} = c_1\mathrm{y}_1 + c_2\mathrm{y}_2 = c_1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \end{aligned}$$

we need to solve the corresponding linear systems with the same coefficient matrix $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. The augmented matrix for all of the three linear systems is $\left[\begin{array}{cc|c|c|c} 1 & -1 & 1 & -1 & 2 \\ 1 & 1 & -1 & 0 & 1 \end{array}\right]$. By converting this into its RREF, we get

$$\left[\begin{array}{cc|c|c|c} 1 & 0 & 0 & -\frac{1}{2} & \frac{3}{2} \\ 0 & 1 & -1 & \frac{1}{2} & -\frac{1}{2} \end{array}\right].$$

Therefore, $(a_1, a_2) = (0, -1)$, $(b_1, b_2) = (-\frac{1}{2}, \frac{1}{2})$, $(c_1, c_2) = (\frac{3}{2}, -\frac{1}{2})$ and hence

$$A' = [T]_\alpha^\beta = [[T(\mathrm{x}_1)]_\beta [T(\mathrm{x}_2)]_\beta [T(\mathrm{x}_3)_\beta]]_{2\times 3} = \begin{bmatrix} 0 & -\frac{1}{2} & \frac{3}{2} \\ -1 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

(Note that $A=[T]=[T]_{\epsilon_1}^{\epsilon_2}=\begin{bmatrix}2&0&-1\\0&1&-1\end{bmatrix}$.) ■

Example 2

Let $T:\mathbb{R}^2\to\mathbb{R}^3$ be defined via $T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\begin{bmatrix}x+y\\x-3y\\-2x+y\end{bmatrix}$ and

$$\alpha=\left\{\mathrm{x}_1=\begin{bmatrix}1\\1\end{bmatrix},\ \mathrm{x}_2=\begin{bmatrix}2\\1\end{bmatrix}\right\},\ \beta=\left\{\mathrm{y}_1=\begin{bmatrix}1\\0\\-1\end{bmatrix},\ \mathrm{y}_2=\begin{bmatrix}-1\\2\\1\end{bmatrix},\ \mathrm{y}_3=\begin{bmatrix}0\\1\\1\end{bmatrix}\right\}$$

be ordered bases for $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. Find $A'=[T]_\alpha^\beta$.

Solution

Since $T(\mathrm{x}_1)=\begin{bmatrix}2\\-2\\-1\end{bmatrix}$, $T(\mathrm{x}_2)=\begin{bmatrix}3\\-1\\-3\end{bmatrix}$, we have

$$[\mathrm{y}_1\ \ \mathrm{y}_2\ \ \mathrm{y}_3 \,\vdots\, T(\mathrm{x}_1) \,\vdots\, T(\mathrm{x}_2)]=\begin{bmatrix}1&-1&0&\vdots&2&\vdots&3\\0&2&1&\vdots&-2&\vdots&-1\\-1&1&1&\vdots&-1&\vdots&-3\end{bmatrix}.$$

We can get its RREF as follows:

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

$$\longrightarrow\begin{bmatrix}1&0&0&\vdots&\frac{1}{2}&\vdots&\frac{5}{2}\\0&1&0&\vdots&-\frac{3}{2}&\vdots&-\frac{1}{2}\\0&0&1&\vdots&1&\vdots&0\end{bmatrix}.$$

Therefore, we get A' as follows:

$$A'=[T]_\alpha^\beta=\begin{bmatrix}\frac{1}{2}&\frac{5}{2}\\-\frac{3}{2}&-\frac{1}{2}\\1&0\end{bmatrix}.$$ □

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

① Write $[\mathrm{y}_1\ \ \mathrm{y}_2\ \ \mathrm{y}_3 \,\vdots\, T(\mathrm{x}_1) \,\vdots\, T(\mathrm{x}_2)]$.

```
x, y = var('x, y')
h(x, y) = [x+y, x-3*y, -2*x+y]
T = linear_transformation(QQ^2, QQ^3, h)
x1=vector([1, 1])
x2=vector([2, 1])
y1=vector([1, 0, -1])
y2=vector([-1, 2, 1])
y3=vector([0, 1, 1])
B=column_matrix([y1, y2, y3, T(x1), T(x2)])   # Matrix whose columns are
                                                 # the vectors defined above
print B
```

```
[ 1 -1  0  2  3]
[ 0  2  1 -2 -1]
[-1  1  1 -1 -3]
```

② RREF $[\mathrm{y}_1 \ \ \mathrm{y}_2 \ \ \mathrm{y}_3 : \ T(\mathrm{x}_1) : \ T(\mathrm{x}_2)]$

```
C=B.echelon_form()
print C
```

```
[   1    0    0  1/2  5/2]
[   0    1    0 -3/2 -1/2]
[   0    0    1    1    0]
```

③ Finding $[T]_{\alpha}^{\beta}$

```
A=C.submatrix(0, 3, 3, 2)  # C.submatrix(a, b, c, d)
# submatrix with c consecutive rows of C starting from row a+1 and d
# consecutive columns of C starting from  column b+1
print A
```

```
[ 1/2  5/2]
[-3/2 -1/2]
[   1    0]
```

We shall include calculation using the inbuilt function. Following are the codes.

```
var('x,y')
```

```
h(x,y)=[x+y,x-3*y,-2*x+y]
V=QQ^2;W=QQ^3
T=linear_transformation(V,W,h)
y1=vector(QQ,[1,1]);y2=vector(QQ,[2,1])
x1=vector(QQ, [1,0,-1]);x2=vector(QQ, [-1,2,1]);x3=vector(QQ,[0,1,1]);
alpha=[y1,y2]; beta=[x1,x2,x3]
V1=V.subspace_with_basis(alpha); W1=W.subspace_with_basis(beta)
T1=(T.restrict_domain(V1)).restrict_codomain(W1)
T1.matrix(side='right')
```

```
[ 1/2  5/2]
[-3/2 -1/2]
[   1    0]
```
■

Example 3

Let $T:\mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation defined as $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)=\begin{bmatrix} 2x+y \\ x-y \\ x+4y \end{bmatrix}$ and consider the ordered bases $\alpha=\{e_2, e_1\}$, $\beta=\{e_3, e_2, e_1\}$ for $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. Answer the following questions:

(1) Find $A'=[T]_\alpha^\beta$.

(2) Compute $\left[T\left(\begin{bmatrix} -4 \\ 6 \end{bmatrix}\right)\right]_\beta$ using $A'=[T]_\alpha^\beta$ in (1).

(3) Using the definition of T, find the standard matrix $A=[T]=[T]_{\epsilon_1}^{\epsilon_2}$ and $T\left(\begin{bmatrix} -4 \\ 6 \end{bmatrix}\right)$, where ϵ_1 and ϵ_2 are the standard bases of $\mathbb{R}^2$ and $\mathbb{R}^3$ repectively.

Solution

(1) Since $T(e_2)=\begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$, $T(e_1)=\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, we get

$\begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}_\beta=\begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}_\beta=\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$. Hence

$$A'=[T]_\alpha^\beta=\begin{bmatrix} 4 & 1 \\ -1 & 1 \\ 1 & 2 \end{bmatrix}.$$

(2) Since $\begin{bmatrix} -4 \\ 6 \end{bmatrix}_\alpha=\begin{bmatrix} 6 \\ -4 \end{bmatrix}$, we have

$$\left[T\left(\begin{bmatrix} -4 \\ 6 \end{bmatrix}\right)\right]_\beta=[T]_\alpha^\beta\begin{bmatrix} -4 \\ 6 \end{bmatrix}_\alpha=A'\left(\begin{bmatrix} -4 \\ 6 \end{bmatrix}\right)_\alpha=\begin{bmatrix} 4 & 1 \\ -1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ -4 \end{bmatrix}=\begin{bmatrix} 20 \\ -10 \\ -2 \end{bmatrix}$$

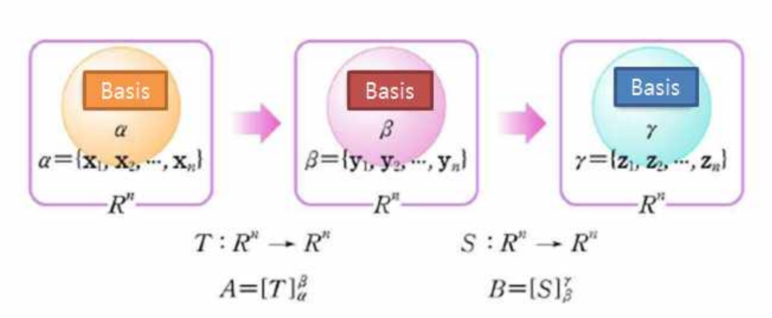
$$=20e_3+(-10)e_2+(-2)\,e_1$$

(Note that $T\left(\begin{bmatrix} -4 \\ 6 \end{bmatrix}\right)=\left[T\left(\begin{bmatrix} -4 \\ 6 \end{bmatrix}\right)\right]_\epsilon=\begin{bmatrix} -2 \\ -10 \\ 20 \end{bmatrix}=(-2)\,e_1+(-10)e_2+20e_3$.)

(3) $[T]_{\epsilon_1}^{\epsilon_2}=\begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 1 & 4 \end{bmatrix}$, $T\left(\begin{bmatrix} -4 \\ 6 \end{bmatrix}\right)=\begin{bmatrix} 2\cdot(-4)+6 \\ -4-6 \\ -4+4\cdot 6 \end{bmatrix}=\begin{bmatrix} -2 \\ -10 \\ 20 \end{bmatrix}$. ■

[Remark] **For given three vector spaces with different ordered bases, we can consider two linear transformations T and S, and their corresponding matrix representations A and B relative to the given ordered bases.**

Composition of Linear Transformations



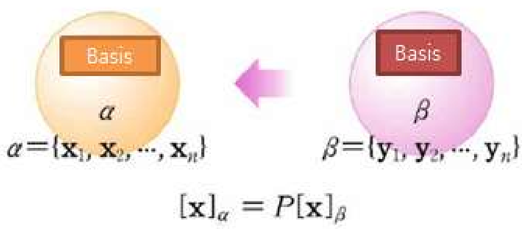
Let T be a linear transformation from a vector space $\mathbb{R}^n$ with an ordered basis α into a vector space $\mathbb{R}^n$ with an ordered basis β, and S be a linear transformation from a vector sapce $\mathbb{R}^n$ with an ordered basis β into a vector space $\mathbb{R}^n$ with an ordered basis γ. Suppose these linear transformations have their corresponding matrix representations $A=[T]_\alpha^\beta$ and $B=[S]_\beta^\gamma$, respectively. We can consider the composition $S\circ T$. Then its matrix representation is

$$[S\circ T]_\alpha^\gamma = [S]_\beta^\gamma [T]_\alpha^\beta = BA,$$

That is, the product of the two matrix representations of T and S.

[Remark] Transition Matrix

As we have discussed earlier, $[\mathrm{x}]_\alpha = [I]_\beta^\alpha [\mathrm{x}]_\beta = P[\mathrm{x}]_\beta$, the matrix $P=[I]_\beta^\alpha$ is called the transition matrix from ordered basis β to ordered basis α. We can consider the transition matrix as linear transformation $T_P(\mathrm{x}) = P\mathrm{x}$.



Example 4

Let $T, S : \mathbb{R}^2 \to \mathbb{R}^2$ be linear transformations defined as

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x+y \\ x-y \end{bmatrix} \text{ and } S\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -x+5y \\ 2x+3y \end{bmatrix}$$

respectively. Consider the ordered bases $\alpha = \{\mathbf{e}_1, \mathbf{e}_2\}$, $\beta = \{\mathbf{e}_2, \mathbf{e}_1\}$, $\gamma = \{(1,0), (1,1)\}$ for $\mathbb{R}^2$. Find the matrix representation of the composition $S \circ T$ with respect to the ordered bases α and γ.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

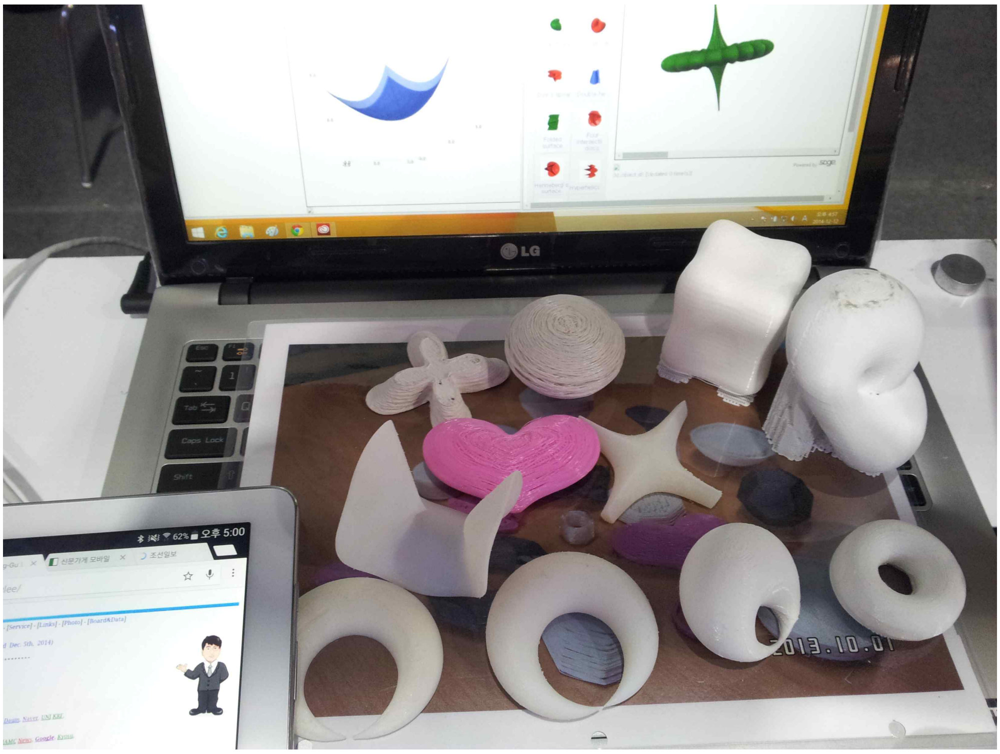
```
x, y = var('x, y')
ht(x, y) = [2*x+y, x-y];hs(x, y) = [-x+5*y, 2*x+3*y]
T = linear_transformation(QQ^2, QQ^2, ht)
S = linear_transformation(QQ^2, QQ^2, hs)
x1=vector([1, 0]);x2=vector([0, 1]);x3=vector([1, 1])
B=column_matrix([x2, x1, T(x1), T(x2)])
C=B.echelon_form()
MT=C.submatrix(0, 2, 2, 2)
print "Matrix of T="
print MT
D=column_matrix([x1, x3, S(x2), S(x1)])
E=D.echelon_form()
MS=E.submatrix(0, 2, 2, 2)
print "Matrix of S="
print MS
print "MS*MT="
print MS*MT
```

```
Matrix of T=
[ 1 -1]
[ 2  1]
Matrix of S=
[ 2 -3]
[ 3  2]
MS*MT=
[-4 -5]
[ 7 -1]
```

```
F=column_matrix([x1, x3, S(T(x1)), S(T(x2))])
G=F.echelon_form()
MST=G.submatrix(0, 2, 2, 2)
```

```
print "Matrix of S*T="
print MST
```

```
Matrix of S*T=
[-4 -5]
[ 7 -1]
```



3D Printing object 1

http://matrix.skku.ac.kr/2014-Album/2014-12-ICT-DIY/index.html

http://youtu.be/FgAzOkqq7Sg

8.2 Similarity and Diagonalization

- Lecture Movie : http://youtu.be/xirjNZ40kRk, http://youtu.be/MnfLcBZsV-I
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-11-sec-8-2.html

In this section, we present various matrix representations of a linear transformation T from $\mathbb{R}^n$ to itself in terms of transition matrix. We also study when the transition matrix becomes a diagonal matrix.

[Remark] **Relationship between matrix representations** $[T]=[T]_{\epsilon_1}^{\epsilon_2}$ **and** $[T]_\alpha^\beta$

$$[T]_\alpha^\beta = [I]_{\epsilon_2}^\beta [T]_{\epsilon_1}^{\epsilon_2} [I]_\alpha^{\epsilon_1} = [I]_{\epsilon_2}^\beta [T] [I]_\alpha^{\epsilon_1}.$$

Theorem 8.2.1

Let $T:\mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and α and β be ordered bases for $\mathbb{R}^n$. If $A=[T]_\alpha$, $A'=[T]_\beta$, then we have

$$A' = P^{-1}AP,$$

where $P=[I]_\beta^\alpha$ is the transition matrix from β to α.

$$\begin{tikzcd}
{[\mathbf{x}]_\beta} \arrow[r, "{[T]_\beta = A'}"] \arrow[d, "\cong", "{P=[I]_\beta^\alpha}"'] & {[T(\mathbf{x})]_\beta} \arrow[d, shift right, "\cong", "P"'] \\
{[\mathbf{x}]_\alpha} \arrow[r, "{A=[T]_\alpha}"'] & {[T(\mathbf{x})]_\alpha} \arrow[u, shift right, "{Q=[I]_\alpha^\beta=P^{-1}}"']
\end{tikzcd}$$

$$A' = [T]_\beta = [I]_\alpha^\beta [T]_\alpha [I]_\beta^\alpha = P^{-1}AP.$$

http://www.math.tamu.edu/~yvorobet/MATH304-503/Lect2-12web.pdf

Example 1

Let $T:\mathbb{R}^2 \to \mathbb{R}^2$ be a linear transfromation defined by $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x-y \\ x+3y \end{bmatrix}$. If α is the standard basis ϵ for $\mathbb{R}^2$ and $\beta = \left\{ \mathrm{y}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ \mathrm{y}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^2$, find $A' = [T]_\beta$ using the transition matrix $P = [I]_\beta^\alpha$.

Solution

Let A be the standard matrix relative to the standard basis $\alpha = \epsilon$ for linear transformation T. Then we can find $A = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}$. If $[I]_\beta^\alpha = P$, then

$$P = [[\mathrm{y}_1]_\alpha \ :\ [\mathrm{y}_2]_\alpha] = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}.$$

Therefore, $P^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$ and by **Theorem 8.2.1** we get A'as follows:

$$A' = P^{-1}AP = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}.$$

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

```
x, y = var('x, y')
h(x, y) = [2*x-y, x+3*y]
T = linear_transformation(QQ^2, QQ^2, h)
x1=vector([1, 0]);x2=vector([0, 1])
y1=vector([0, 1]);y2=vector([-1, 1])
B=column_matrix([x1, x2, y1, y2])
C=B.echelon_form()
P=C.submatrix(0, 2, 2, 2)
print "Transition Matrix="
print P
A = T.matrix(side='right')
print "A="
print A
print "P.inverse()*A*P"
print P.inverse()*A*P
D=column_matrix([y1, y2, T(y1), T(y2)])
E=D.echelon_form()
print "Matrix of A wrt beta="
print E.submatrix(0, 2, 2, 2)
```

Transition Matrix=
[0 -1]
[1 1]
A=
[2 -1]
[1 3]
P.inverse()*A*P
[2 -1]
[1 3]
Matrix of A wrt beta=
[2 -1]
[1 3] ■

Similarity

Definition [Similarity]

For square matrices A, B of the same order, if there exists an invertible matrix P such that

$$B = P^{-1}AP,$$

then we say that B is similar to A. We use $B \sim A$ for similar matrices A, B.

Example 2

For $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, it can be shown that $B = P^{-1}AP$. Hence B is similar to A, which is denoted by $B \sim A$. ■

Theorem 8.2.2

For square matrices A, B, C of the same order, the following hold:

(1) $A \sim A$

(2) $B \sim A \Rightarrow A \sim B$

(3) $B \sim A,\ A \sim C \Rightarrow B \sim C$

Therefore, the similarity relation is an equivalence relation.

Theorem 8.2.3

For square matrices A, B of the same order, if A, B are similar to each other, then we have the following:

(1) $\det(A) = \det(B)$.

(2) $\mathrm{tr}(A) = \mathrm{tr}(B)$.

Proof Since $A \sim B$, there exists an invertible matrix P such that $A = P^{-1}BP$.

(1) By the multiplicative property of determinant,

$$\det(A) = \det(P^{-1}BP)$$

$$\Rightarrow \det(A) = \det(P^{-1})\det(B)\det(P) \quad (\because \det(AB) = \det(A)\det(B))$$

$$\Rightarrow \det(A) = \det(P^{-1})\det(P)\det(B)$$

$$\Rightarrow \det(A) = \det(B) \quad (\because \det(P^{-1}P) = 1 = \det(P^{-1})\det(P))$$

(2) $\mathrm{tr}(A) = \mathrm{tr}(P^{-1}BP) = \mathrm{tr}(BPP^{-1}) \quad (\because \mathrm{tr}(AS) = \mathrm{tr}(SA))$

$= \mathrm{tr}(BI) = \mathrm{tr}(B)$ ■

Since similar matrices have the same determinant, it follows that they have the same characteristic equation and hence the same eigenvalues. For a square matrix, in ordered to solve problems of determinant and/or eigenvalues, we can use its similar matrices which make the problems simpler.

Diagonalizable Matrices

Definition [Diagonalizable Matrices]

Suppose a square matrix A is simlar to a diagonal matrix, that is, there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix. Then A is called **diagonalizable** and the invertible matrix P is called a diagonalizing matrix for A.

- If $P^{-1}AP = D$, then $A = PDP^{-1}$ and hence we have

$$\begin{aligned} A^k &= (PDP^{-1})^k = (PDP^{-1})(PDP^{-1}) \cdots (PDP^{-1}) \quad (k \text{ multiplications of } PDP^{-1}) \\ &= PD(P^{-1}P)D(P^{-1}P) \cdots (P^{-1}P)DP^{-1} \\ &= PD^kP^{-1} \end{aligned}$$

This implies that if a matrix is diagonalizable, then its powers can be very easily computed.

Example 3

For invertible matrix $P = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$ and matrix $A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$, we have

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

Hence A is diagonalizable. □

- http://matrix.skku.ac.kr/RPG_English/8-TF-diagonalizable.html

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

```
A=matrix(QQ, [[1, 1], [-2, 4]])
print A.is_diagonalizable()          # Checking if diagonalizable
```

```
True
```
■

Example 4

Since every diagonal matrix D satisfies $I^{-1}DI = D$, it is diagonalizable. ■

Example 5 Show that $A=\begin{bmatrix}0&1\\0&0\end{bmatrix}$ is not diagonalizable.

Solution

Suppose to the contrary that A is diagonalizable, that is, there exist an invertible matrix P and a diagonal matrix D with

$$P=\begin{bmatrix}a&b\\c&d\end{bmatrix},\ (ad-bc\neq 0),\ D=\begin{bmatrix}e&0\\0&f\end{bmatrix},$$

such that $P^{-1}AP=D$. Since $AP=PD$, we have

$$\begin{bmatrix}a&b\\c&d\end{bmatrix}\begin{bmatrix}e&0\\0&f\end{bmatrix}=\begin{bmatrix}0&1\\0&0\end{bmatrix}\begin{bmatrix}a&b\\c&d\end{bmatrix},$$

which gives $\begin{bmatrix}ae&bf\\ce&df\end{bmatrix}=\begin{bmatrix}c&d\\0&0\end{bmatrix}$. Hence $ce=0$.

If $c\neq 0$, then $e=0$ and $ae=0=c(\neq 0)$. Hence $c=0$. Similarly, we can show that $d=0$. The conditions $c=0$ and $d=0$ give a contradiction to $ad-bc\neq 0$. Therefore, A is not diagonalizable. ■

Equivalent Condition for Diagonalizability

Theorem 8.2.4 [Equivalent Condition]

Let A be a square matrix of order n. Then A is diagonalizable if and only if A has n linearly independent eigenvectors. Furthermore, if A is diagonalizable, then A is similar to diagonal matrix D whose main diagonal entries are equal to the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A, and the ith column of a diagonalizing matrix P is an eigenvector of A corresponding to eigenvalue λ_i.

Proof ($\Rightarrow$) If A is diagonalizable, then there exists an invertible matrix

$$P=\left[\mathbf{p}^{(1)} : \mathbf{p}^{(2)} : \cdots : \mathbf{p}^{(n)}\right]$$

such that $P^{-1}AP = B$ where $B = \mathrm{diag}(b_1, b_2, \cdots, b_n)$. Since $AP = PB$, we get $A\mathbf{p}^{(1)} = b_1\mathbf{p}^{(1)}$, $A\mathbf{p}^{(2)} = b_2\mathbf{p}^{(2)}$, ... , $A\mathbf{p}^{(n)} = b_n\mathbf{p}^{(n)}$. Hence $b_1, b_2, \ldots, b_n$ are eigenvalues of A and $B = D$. Note that $\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \ldots, \mathbf{p}^{(n)}$ are eigenvectors corresponding to $b_1 = \lambda_1, b_2 = \lambda_2, \ldots, b_n = \lambda_n$, respectively. Since P is invertible, it follows that its columns $\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \ldots, \mathbf{p}^{(n)}$ are linearly independent.

($\Leftarrow$) Suppose A has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and their corresponding eigenvectors $\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \ldots, \mathbf{p}^{(n)}$ that are linearly independent. Then we can construct a matrix P as follows:

$$P = \left[\mathbf{p}^{(1)} : \mathbf{p}^{(2)} : \ldots : \mathbf{p}^{(n)}\right].$$

Then

$$AP = \left[A\mathbf{p}^{(1)} : A\mathbf{p}^{(2)} : \ldots : A\mathbf{p}^{(n)}\right] = \left[\lambda_1\mathbf{p}^{(1)} : \lambda_2\mathbf{p}^{(2)} : \cdots : \lambda_n\mathbf{p}^{(n)}\right]$$
$$= \left[\mathbf{p}^{(1)} : \mathbf{p}^{(2)} : \cdots : \mathbf{p}^{(n)}\right]\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = PD.$$

Since the columns of P are linearly independent, the matrix P is invertible, giving $P^{-1}AP = D$. Therefore A is diagonalizable. ■

[Remark] **Procedure for diagonalizing a matrix A**

- Step 1: Find n linearly independent eigenvectors $\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \ldots, \mathbf{p}^{(n)}$ of A.
- Step 2: Construct a matrix P whose columns are $\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \ldots, \mathbf{p}^{(n)}$ in this order.
- Step 3: The matrix P diagonalizes A and $P^{-1}AP$ is a diagonal matrix whose main diagonal entries are eigenvalues $\lambda_1, \ldots, \lambda_n$ of A

$$D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Example 6

It can be shown that the matrix $A = \begin{bmatrix} 5 & -6 \\ 2 & -2 \end{bmatrix}$ has eigenvalues $\lambda_1 = 2, \lambda_2 = 1$ and their corresponding eigenvectors are

$\mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, respectively. Since these eigenvectors are linearly independent, by Theorem **8.2.4,** A is diagonalizable. If $P = [\mathbf{x}_1 \ \mathbf{x}_2] = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$, then we have

$$P^{-1}AP = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 5 & -6 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$ ■

Example 7

Show that $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$ is diagonalizable and find the diagonalizing matrix P of A.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

```
A=matrix([[0, 0, -2], [1, 2, 1], [1, 0, 3]])
print A.eigenvalues()            # Eigenvalue Computation
```

```
[1, 2, 2]
```

A has eigenvalues $\lambda_1 = 1, \lambda_2 = 2$. We now compute linearly independent eigenvectors of A.

For $\lambda_1 = 1$, we solve $A\mathbf{x} = \lambda_1 \mathbf{x}$ (that is, $(\lambda_1 I - A)\mathbf{x} = \mathbf{0}$) for $\mathbf{x}$.

```
E=identity_matrix(3)
print (E-A).echelon_form()
```

```
[ 1  0  2]
[ 0  1 -1]
[ 0  0  0]
```

Since $\mathbf{x} = \begin{bmatrix} -2t \\ t \\ t \end{bmatrix} = t \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$ $(t \in R)$, we get $\mathbf{x}_1 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$; For $\lambda_2 = 2$, we solve $A\mathbf{x} = \lambda_2 \mathbf{x}$ (that is, $(\lambda_2 I - A)\mathbf{x} = \mathbf{0}$) for $\mathbf{x}$.

```
print (2*E-A).echelon_form()
```

```
[1 0 1]
[0 0 0]
[0 0 0]
```

This gives $\mathbf{x} = \begin{bmatrix} -s \\ t \\ s \end{bmatrix} = s\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + t\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ $(s, t \in R)$ and hence

$$\mathbf{x}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix},\ \mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

```
x1=vector([-2, 1, 1])
x2=vector([-1, 0, 1])
x3=vector([0, 1, 0])
P=column_matrix([x1, x2, x3])
print P
print
print P.det()
```

```
[-2 -1  0]
[ 1  0  1]
[ 1  1  0]

1
```

Since the above computation shows that the determinant of P is not zero, P is invertible. Hence its columns $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ are linearly independent. Therefore, by Theorem **8.2.4**, A is diagonalizable.

```
print P^-1*A*P # Computing diagonal matrix whose main diagonal entries
               # are eigenvalues of A.
```

```
[1 0 0]
[0 2 0]
[0 0 2]
```

■

Theorem 8.2.5

If $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ are eigenvectors of $A = [a_{ij}]_{n \times n}$ corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$, then the set $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ is linearly independent.

Proof (Exercise) Hint. This can be proved by the mathematical induction k.

Theorem 8.2.6

If a square matrix A of order n has n distinct eigenvalues, then A is diagonalizable.

Proof Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ be eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then, by Theorem **8.2.5**, the eigenvectors are linearly independent. Therefore, Theorem **8.2.4** implies that A is diagonalizable. ■

Example 8 The matrix $A = \begin{bmatrix} 5 & -6 \\ 2 & -2 \end{bmatrix}$ in Example 6 has two distinct eigenvalues. Thus, by Theorem **8.2.6**, A is diagonalizable. ■

Note that a diagonal matrix A can have a repeated eigenvalue. Therefore, the converse of Theorem **8.2.6** is not necessarily true.

Algebraic Multiplicity and Geometric Multiplicity of an Eigenvalue

Definition [Algebraic and Geometric Multiplicity]

Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be distinct eigenvalues of $A = [a_{ij}]_{n \times n}$. Then the characteristic polynomial of A can be written as

$$|\lambda I - A| = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_k)^{m_k}$$

Algebraic Multiplicity

In the above expression the sum of the exponents $m_1, m_2, \ldots, m_k$ is equal to n. The positive integer m_i is called the algebraic multiplicity of λ_i and the number of linearly independent eigenvectors corresponding to the eigenvalue λ_i is called the geometric multiplicity of λ_i.

Theorem 8.2.7 [Equivalent Condition for Diagonalizability]

Let A be a square matrix of order n. Then A is diagonalizable if and only if the sum of the geometric multiplicities of eigenvalues of A is equal to n.

Proof By **Theorem 8.2.4** an equivalent condition for a square matrix A of order n to diagonalizable is to have n linearly independent eigenvectors. Since the sum of the geometric multiplicities of eigenvalues of A is equal to the number of linearly independent eigenvectors of A and it is equal to n, the result follows. ■

Theorem 8.2.8

Let A be a square matrix and λ be an eigenvalue of A. Then the algebraic multiplicity of λ is greater than or equal to the geometric multiplicity of λ.

Proof Let k be the geometric multiplicity of an eigenvalue λ of A, and let P_1 be the $n \times k$ matrix whose columns are the k linearly independent eigenvectors of A corresponding to eigenvalue λ. We can construct an invertible matrix P by adding $n-k$ linearly independent columns to P_1. Let $P = [P_1\, P_2]$ be the resulting invertible matrix and let $Q = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}$ be the inverse of P. Then $QAP = P^{-1}AP = \begin{bmatrix} \lambda I_k & * \\ O & * \end{bmatrix}$. Note that A and $P^{-1}AP$ have same characteristic polynomials. Since $P^{-1}AP$ has first k columns have λ in its diagonal, the characteristic polynomial of $P^{-1}AP$ has a factor of at least $(x-\lambda)^k$. Hence, the algebraic multiplicity of λ is greater than or equal to the geometric multiplicity of λ. ■

Theorem 8.2.9 [Equivalent Condition for Diagonalizability]

Let A be a square matrix of order n. Then A is diagonalizable if and only if each eigenvalue λ of A has the same algebraic and geometric multiplicity.

Proof If A is diagonalizable, then there exists an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$, or equivalently $AP = PD$. This implies that A times column i of P is equal to scalar multiple of the column i of P.

Hence, all the n columns of P are eigenvectors of A, which implies that each eigenvalue of A has the same algebraic and geometric multiplicity. The converse is also true by Theorem 8.2.5. ■

Example 9

For $A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ -3 & 5 & 3 \end{bmatrix}$, its characteristic equation is $|\lambda I_2 - A| = (\lambda-2)(\lambda-3)^2 = 0$. Hence the eigenvalues of A are $\lambda = 2,\ 3$ and $\lambda = 3$ has algebraic multiplicity 2. The following two vectors are linearly independent eigenvectors of A

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \\ 8 \end{bmatrix},\ \mathbf{x}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

However, matrix A cannot have three linearly independent eigenvectors and hence Theorem **8.2.4** implies that A is not diagonalizable. ■

Example 10

It can be shown that $A = \begin{bmatrix} 4 & -2 & 1 \\ 2 & 0 & 1 \\ 2 & -2 & 3 \end{bmatrix}$ has eigenvalues 3 and 2 with algebraic multicity 1 and 2 respectively, We can further show that geometric multiplicity of 3 and 2 are 1 and 2 respectively. Hence A is diagonalizable. It can be verified $P = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}$ diagonalizes A and $A = PDP^{-1}$, where $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. Let us further compute A^5.

$$A^5 = PD^5P^{-1} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2^5 & 0 & 0 \\ 0 & 2^5 & 0 \\ 0 & 0 & 3^5 \end{bmatrix} \begin{bmatrix} -2 & 3 & -1 \\ -1 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 454 & -422 & 211 \\ 422 & -390 & 211 \\ 422 & -422 & 243 \end{bmatrix}$$ ■

8.3 Diagonalization with orthogonal matrix, *Function of matrix

- Lecture Movie : http://youtu.be/jimlkBGAZfQ, http://youtu.be/B--ABwoKAN4
- ab : http://matrix.skku.ac.kr/knou-knowls/cla-week-11-sec-8-3.html

Symmetric matrices appear in many applications. In this section, we study useful properties of symmetric matrices and show that every symmetric matrix is orthogonally diagonalizable. Furthermore, we study matrix functions using matrix diagonalization.

Orthogonal Matrix

Definition [Orthogonal Matrix]

For real square matrix A, if A is invertible and $A^{-1} = A^T$, then A is called an **orthogonal matrix**.

Theorem 8.3.1

If A is an orthogonal matrix of order n, then the following hold:

(1) The rows of A are unit vectors and they are perpendicular to each other.
(2) The columns of A are unit vectors and they are perpendicular to each other.
(3) A is invertible.
(4) $\|A\mathbf{x}\| = \|\mathbf{x}\|$ for any $n \times 1$ vector $\mathbf{x}$ (Norm Preserving Property).

Proof Similar to the proof of Theorem **6.2.3**.

Example 1

For $A = \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix}$, we have $A^T = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix}$. Since

$$A^TA = \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix} = I_3, \quad A \text{ is an orthogonal matrix.} \quad \blacksquare$$

- The inverse of an orthogonal matrix can be obtained by taking transposition of the orthogonal matrix.

Definition **[Orthogonal Similarity]**

Let A and C be square matrices of the same order. If there exists an orthogonal matrix P such that $C = P^TAP$, then C is said to be orthogonally similar to A.

Definition **[Orthogonally Diagonalizable]**

For a square matrix A, if there exists an orthogonal matrix diagonalizing A, then A is called orthogonally diagonalizable and P is called a matrix orthogonally diagonalizing A.

What matrices are orthogonally diagonalizable? (Symmetric Matrices)

Theorem 8.3.2

Every eigenvalue of a real symmtric matrix is a real number.

Proof (Exercise) http://www.quandt.com/papers/basicmatrixtheorems.pdf

Example 2 The symmetric matrix $A = \begin{bmatrix} 0 & 0 & -2 \\ 0 & -2 & 0 \\ -2 & 0 & 3 \end{bmatrix}$ has characteristic equation $|\lambda I_3 - A| = (\lambda+2)(\lambda-4)(\lambda+1) = 0$ and hence its eigenvalues are

$\lambda_1 = -2,\ \lambda_2 = 4,\ \lambda_3 = -1$ that are all real numbers. ■

Theorem 8.3.3

If a square matrix A is symmetric, then eigenvectors of A corresponding to distinct eigenvalues are perperdicular to each other.

Proof (Exercise) http://www.quandt.com/papers/basicmatrixtheorems.pdf

Theorem 8.3.4

Let A be a square matrix. Then A is orthogonally diagonalizable if and only if the matrix A is symmetric.

Proof ($\Rightarrow$) Suppose A is orthogonally diagonalizable. Then there exist an orthogonal matrix P and a diagonal matrix D such that $P^TAP = D$. Since $D = D^T$, we have

$$P^TAP = D = D^T = (P^TAP)^T = P^TA^TP.$$

Hence

$$\begin{aligned} P^TAP = P^TA^TP \quad &\Leftrightarrow \quad P(P^TAP)P^T = P(P^TA^TP)P^T \\ &\Leftrightarrow \quad (PP^T)A(PP^T) = (PP^T)A^T(PP^T) \\ &\Leftrightarrow \quad A = A^T \end{aligned}$$

Therefore, A is symmetric.

($\Leftarrow$) : Exercise ■

Theorem 8.3.5

If A is a symmetric matrix of order n, then A has n eigenvectors forming an orthonormal set.

Proof Since A is symmetric, by Theorem **8.3.4**, A is orthogonally diagonalizabl, that is, there exist an orthogonal matrix P and a diagonal matrix D such that $P^TAP = D$. Hence the main diagonal entries of D are the eigenvalues of A and the columns of P are n eigenvectors of A. Since the columns of the orthogonal matrix P form an orthognormal set, the n eigenvectors of A are orthonormal. ■

Theorem 8.3.6

For a square matrix A of order n, the following are equivalent:

(1) A is orthogonally diagonalizable.
(2) A has n eigenvectors that are orthonormal.
(3) A is symmetric.

How to find an orthogonal matrix P diagonalizing a given symmetric matrix A?

Example 3

For symmetric matrix $A=\begin{bmatrix}-1 & -1 & 1\\ -1 & 2 & 4\\ 1 & 4 & 2\end{bmatrix}$, find an orthogonal matrix P diagonalizing A.

Solution

Since the characteristic equation of A is $|\lambda I_3 - A| = \lambda(\lambda+3)(\lambda-6)=0$, the eigenvalues of A are $\lambda_1 = -3, \lambda_2 = 0$, $\lambda_3 = 6$. Note that all the eigenvalues are distinct. Hence there exist eigenvectors of A that are orthogonal:

$$\mathbf{x}_1 = \begin{bmatrix}-1\\ -1\\ 1\end{bmatrix},\ \mathbf{x}_2 = \begin{bmatrix}2\\ -1\\ 1\end{bmatrix},\ \mathbf{x}_3 = \begin{bmatrix}0\\ 1\\ 1\end{bmatrix}.$$

By normalizing $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, we get an orthogonal matrix P diagonalizing A:

$$\left\{\begin{bmatrix}-\frac{1}{\sqrt{3}}\\ -\frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{3}}\end{bmatrix}, \begin{bmatrix}\frac{2}{\sqrt{6}}\\ -\frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{6}}\end{bmatrix}, \begin{bmatrix}0\\ \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}\right\} \quad \therefore\ P = \begin{bmatrix}-\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0\\ -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}}\end{bmatrix}$$ ■

Example 4

It can be shown that the matrix $A = \begin{bmatrix} 0 & 3 & 3 \\ 3 & 0 & 3 \\ 3 & 3 & 0 \end{bmatrix}$ has eigenvalues $\lambda_1 = \lambda_2 = -3$ (algebraic multiplicity 2) and $\lambda_3 = 6$. Hence we need to check if $\lambda_1 = -3$ has two linearly independent eigenvectors. After eigenvector computation, we get

$$\mathbf{x}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{x}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

that are linearly independent eigenvectors corresponding to eigenvalue -3. Using the Gram-Schmidt Orthonormalization, we get

$$\mathbf{y}_1 = \mathbf{x}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{y}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \ \mathbf{y}_1}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$$

$$\mathbf{z}_1 = \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \ \mathbf{z}_2 = \frac{\mathbf{y}_2}{\|\mathbf{y}_2\|} = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}.$$

We can find an eigenvector $\mathbf{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ corresponding to the eigenvalue $\lambda_3 = 6$ and normalization gives

$$\mathbf{z}_3 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}.$$

Therefore, the orthogonal matrix P diagonalizing A is given by

$$P = [\mathbf{z}_1 : \mathbf{z}_2 : \mathbf{z}_3] = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}. \qquad \blacksquare$$

[Remark] $*$ **Function of Matrices**

- There are several techniques for lifting a real function to a square matrix function such that interesting properties are maintained. You can read the details in the following:
- https://en.wikipedia.org/wiki/Matrix_function
- http://youtu.be/B--ABwoKAN4

$$e^{A} = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

$$e^{At} = I + At + \frac{A^2t^2}{2!} + \frac{A^3t^3}{3!} + \cdots$$

$$f\left(\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \vdots & \vdots \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \cdots & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix}\right) = \begin{bmatrix} \frac{f(\lambda)}{0!} & \frac{f'(\lambda)}{1!} & \frac{f''(\lambda)}{2!} & \cdots & \frac{f^{(n)}(\lambda)}{n!} \\ 0 & \frac{f(\lambda)}{0!} & \frac{f'(\lambda)}{1!} & \vdots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \cdots & \ddots & \frac{f(\lambda)}{0!} & \frac{f'(\lambda)}{1!} \\ 0 & \cdots & \cdots & 0 & \frac{f(\lambda)}{0!} \end{bmatrix}.$$

[Automobiles with polygonal wheels and the roads customized to the polygonal wheels]

8.4 Quadratic forms

- Lecture Movie : http://youtu.be/vWzHWEhAd-k, http://youtu.be/lznsULrqJ_0
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-12-sec-8-4.html

A quadratic form is a polynomial each of whose terms is quadratic. Quadratic forms appear in many scientific areas including mathematics, physics, economics, statistics, and image processing. Symmetric matrices play an important role in analyzing quadratic forms. In this section, we study how diagonalization of symmetric matrices can be applied to analyse quadratic forms.

Definition

An implicit equation in variables x, y for a quadratic curve can expressed as

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0. \qquad (1)$$

.

This can be rewritten in matrix-vector form as follows:

$$[x,\ y]\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + [d,\ e]\begin{bmatrix} x \\ y \end{bmatrix} + f = 0. \qquad (2)$$

[Remark] **Graph for a quadratic curve (conic section)**

The following are the types of conic sections:

① Non-degenerate conic sections: Circle, Ellipse, Parabola, Hyperbola. See Figure 1.

② Imaginary conic section: There are no points $(x,\ y) \in R^2$ satisfying (1)

③ Degenerate conic section: The graph of the equation (1) is one point, one line, a pair of lines, or having no points.

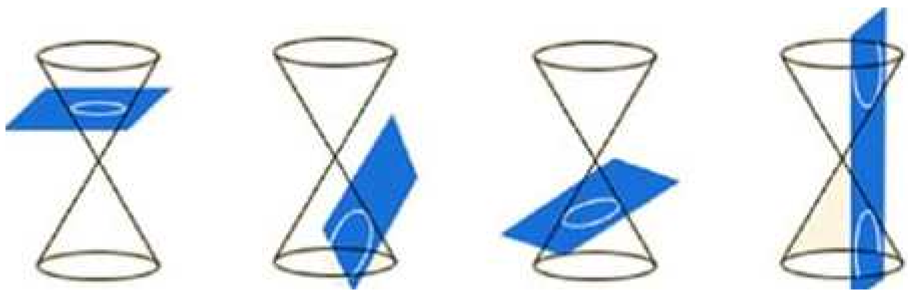
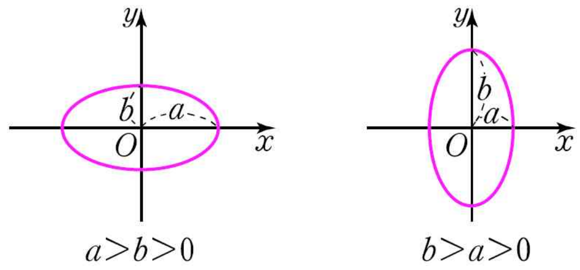


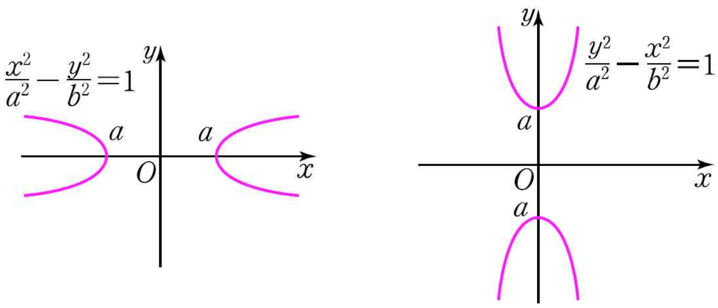
Figure 1

[Remark] Conic Sections in the Standard Position

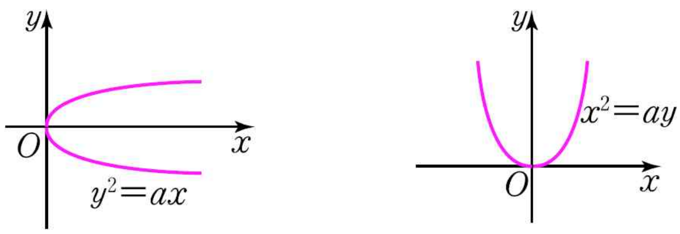
- $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ (Ellipse) (3)



- $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=1$ or $\dfrac{y^2}{a^2}-\dfrac{x^2}{b^2}=1$ (Hyperbola) (4)



- $y^2=ax$ or $x^2=ay$ (Parabola, $a>0$) (5)

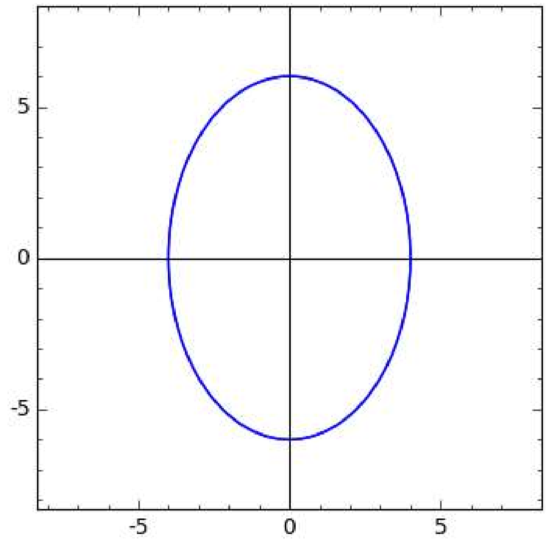


Example 1 (Non-degenerate conic section)

Since the equation $9x^2+4y^2-144=0$ can be written as $\dfrac{x^2}{16}+\dfrac{y^2}{9}=1$, the graph of this equation is an ellipse. The equation $9x^2-4y^2+144=0$ has the standard form $\dfrac{y^2}{9}-\dfrac{x^2}{16}=1$ and hence its graph is a hyperbola. Since the equation $y^2+3x=0$ can be put into $y^2=-3x$, its graph is a parabola.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

```
var('x,y')
implicit_plot(9*x^2+4*y^2-144==0,(x,-8,8),(y,-8,8),axes=True,figsize=5)
```



■

Example 2 (Degenerate conic section)

The graph of the equation $x^2=0$ is the y-axis. The graph of $y^2-9=0$ consists of the two horizontal lines $y=3$, $y=-3$. The graph of $x^2-y^2=0$ consists of the two lines $y=x$ and $y=-x$. The graph of $x^2+y^2=0$ consists of one point $(0,0)$. The graph of $x^2+y^2+1=0$ has no points. ■

- The graph of a quadratic equations with both x^2 and x terms or both y^2 and y terms is a translation of a conic section in the standard position.

Example 3 Let us plot the graph of $3x^2-2y^2-18x+4y+19=0$. By completing squares in $3x^2-2y^2-18x+4y+19=0$, we get

$$3(x-3)^2-2(y-1)^2=6\ . \qquad (6)$$

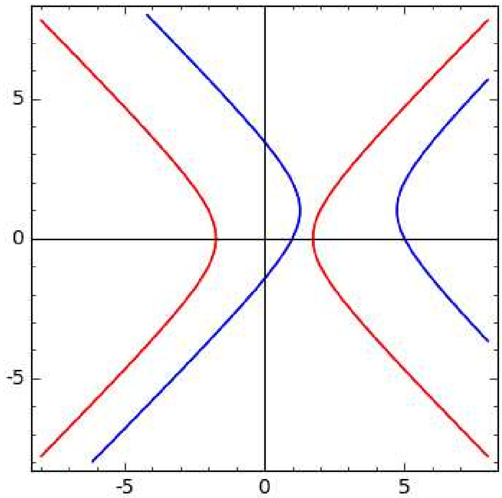
Hence by using the substitutions $x'=x-3$, $y'=y-1$, we get

$$\frac{(x')^2}{2}-\frac{(y')^2}{3}=1$$

in the $x'y'$-coordinate plane. This equation gives a hyperbola of the standard position in the $x'y'$-coordinate plane. Hence the graph of the equation (6) is obtained by translating the hyperbola in the standard position 3 units along the x-axis and 1 unit along the y-axis.

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

```
var('x,y')
c1=implicit_plot(x^2/3-y^2/3-1==0,(x,-8,8),(y,-8,8),axes=True,figsize=5,color='red')
c2=implicit_plot((x-3)^2/3-(y-1)^2/3-1==0,(x,-8,8),(y,-8,8),axes=True,figsize=5,color='blue')
c1+c2
```



■

Quadratic Form

Definition [Quadratic Form]

$$[x,\ y]\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2 \qquad (7)$$

is called the quadratic form of the quadratic equation (1).

Example 4 The quadratic equations $2x^2+6xy+y^2$, x^2+y^2 are quadratic forms, but the quadratic equation $3x^2-6xy+y^2-3x+1$ has a linear term $-3x$ and constant term 1 and hence it is not a quadratic form. ■

A quadratic form can be written in the form of $\mathbf{x}^T A\mathbf{x}$. For example,

$$3x^2+7y^2-2xy=[x\ \ y]\begin{bmatrix}3 & -1\\ -1 & 7\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} \text{ or } 3x^2+7y^2-2xy=[x\ \ y]\begin{bmatrix}3 & -2\\ 0 & 7\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}.$$

This means that the matrix A above is not unique.

- We will use a symmetric matrix A to write a quadratic form:

$$ax^2+by^2+cxy=[x\ \ y]\begin{bmatrix}a & \frac{c}{2}\\ \frac{c}{2} & b\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$$

$$ax^2+by^2+cz^2+dxy+exz+fyz=[x\ \ y\ \ z]\begin{bmatrix}a & \frac{d}{2} & \frac{e}{2}\\ \frac{d}{2} & b & \frac{f}{2}\\ \frac{e}{2} & \frac{f}{2} & c\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}.$$

We use symmetric matrices to represent quadratic forms because symmetric matrices are orthogonally diagonalizable.

Definition

Let $A=[a_{ij}]$ be a symmetric matrix of order n and $\mathbf{x}=\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix}$ for n real values $x_1, x_2, \ldots, x_n$. Then $q(\mathbf{x})=\langle A\mathbf{x},\mathbf{x}\rangle=\mathbf{x}^T A\mathbf{x}=\sum_{i,j=1}^{n} a_{ij}x_i x_j$ is called a quadratic form in $\mathbb{R}^n$.

For a quadratic form in x and y, the xy-term is called a cross-product term. Using orthogonal diagonalization, we can eliminate the cross-product term.

● For a quadratic form

$$q(\mathrm{x}) = \mathrm{x}^T A \mathrm{x} = ax^2 + 2bxy + cy^2,$$

the matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is symmetric, we can find orthonormal eigenvectors $\mathrm{v}_1, \mathrm{v}_2$ corresponding to the eigenvalues λ_1, λ_2 of A. The matrix $P = [\mathrm{v}_1 \ \mathrm{v}_2]$ is orthogonal and P orthogonally diagonalizes A, that is, $P^TAP = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$. Since we can switch the roles of v_1 and v_2 by switching the roles of λ_1 and λ_2, without loss of generality, we can assume $\det(P) = 1$.

Therefore, we can consider P as the rotation matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ in R^2. Let $\mathrm{x} = P\mathrm{x}'$ for some $\mathrm{x}' = \begin{bmatrix} x' \\ y' \end{bmatrix}$. Then

$$\begin{aligned} q(\mathrm{x}) &= \mathrm{x}^T A \mathrm{x} = (P\mathrm{x}')^T A (P\mathrm{x}') = (\mathrm{x}')^T (P^TAP)\mathrm{x}' \\ &= [x' \ y'] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = \lambda_1 (x')^2 + \lambda_2 (y')^2 \end{aligned}$$

and hence q is a quadratic form without any cross-product term in the $x'y'$-coordinate system. Therefore, we get the following theorem.

Theorem 8.4.1 [Diagonalization of a Quadratic Form]

Suppose a symmetric matrix $A = [a_{ij}]_{2\times 2}$ has λ_1, λ_2 as its eigenvalues. Then, by rotating the coordinate axes, the quadratic form $q(\mathrm{x}) = \mathrm{x}^T A \mathrm{x}$ can be written as follows in the $x'y'$-coordinate system

$$q(\mathrm{x}) = \lambda_1 (x')^2 + \lambda_2 (y')^2 . \qquad (8)$$

If the determinant of P is 1 and P diagonalizes A, then the rotation can be obtained by $P^T\mathrm{x} = \mathrm{x}'$ or $\mathrm{x} = P\mathrm{x}'$.

Example 5 Using diagonalization of a quadratic form, determine which conic section the following quadratic equation describes.

$$3x^2 + 2xy + 3y^2 - 8 = 0. \tag{9}$$

Solution

The quadratic equation $3x^2 + 2xy + 3y^2 - 8 = 0$ can be written as

$$\mathbf{x}^T A\mathbf{x} = [x \ \ y]\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 8.$$

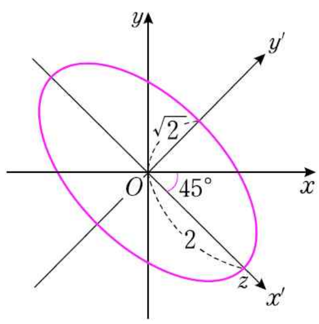
Since the characteristic equation of the symmetric matrix A is $|A - \lambda I| = (3-\lambda)^2 - 1 = (\lambda - 2)(\lambda - 4) = 0$, the eigenvalues of A are $\lambda_1 = 2, \lambda_2 = 4$. By Theorem **8.4.1**, $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = 2(x')^2 + 4(y')^2$. Hence, in the new coordinate system, the quadratic equation becomes

$$2(x')^2 + 4(y')^2 = 8.$$

Since eigenvectors corresponding to $\lambda_1 = 2, \lambda_2 = 4$ are

$$\mathbf{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix},$$

respectively, the orthogonal matrix P diagonalizing A is



$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \cos(-45°) & -\sin(-45°) \\ \sin(-45°) & \cos(-45°) \end{bmatrix}.$$

Therefore $x'y'$-coordinate axes are obtained by rotating the xy-axis 45° clockwise and the equation (9) is an ellipse in the standard position relative to the $x'y'$-coordinate system. ■

Example 6

Sketch the graph of the following quadratic equation

$$34x^2 - 24xy + 41y^2 - 40x - 30y - 25 = 0. \tag{10}$$

Solution

Letting $A = \begin{bmatrix} 34 & -12 \\ -12 & 41 \end{bmatrix}$, $B = [-40, \ -30]$, $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, we can rewrite the equation (10) as follows:

$$\mathbf{x}^T A\mathbf{x} + B\mathbf{x} - 25 = 0 . \qquad (11)$$

Using rotation we first eliminate the cross-product terms. Since the characteristic equation of A is

$$|A - \lambda I| = (\lambda - 25)(\lambda - 50) = 0,$$

the eigenvalues of A are $\lambda_1 = 25$, $\lambda_2 = 50$ and their corresponding orthonormal eigenvectors are $\mathbf{v}_1 = \frac{1}{5}\begin{bmatrix} 4 \\ 3 \end{bmatrix}$, $\mathbf{v}_2 = \frac{1}{5}\begin{bmatrix} -3 \\ 4 \end{bmatrix}$, respectively. Hence we can take $P = [\mathbf{v}_1 : \mathbf{v}_2] = \frac{1}{5}\begin{bmatrix} 4 & -3 \\ 3 & 4 \end{bmatrix}$.

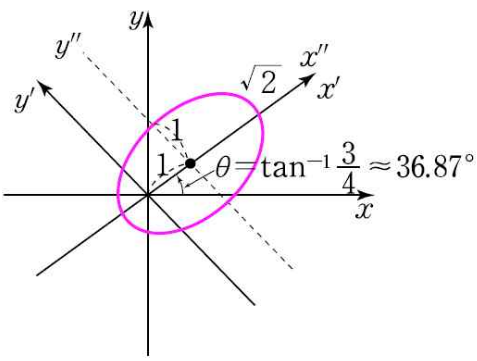
Using axis rotation $\mathbf{x} = P\mathbf{x}'$, we get $\mathbf{x}^T A\mathbf{x} = 25(x')^2 + 50(y')^2$ and $B\mathbf{x} = BP\mathbf{x}' = -50x'$ and hence from (11) we obtain

$$25(x')^2 + 50(y')^2 - 50x' - 25 = 0. \qquad (12)$$

We now use horizontal translation to remove x'-term in (12). By completing the squares in (12) we get

$$25[(x')^2 - 2(x') + 1] + 50(y')^2 = 25 + 25 = 50.$$

That is, $25(x'-1)^2 + 50(y')^2 = 50$. Therefore, the equation (12) repesents an ellipse in the $x''y''$-coordinate system

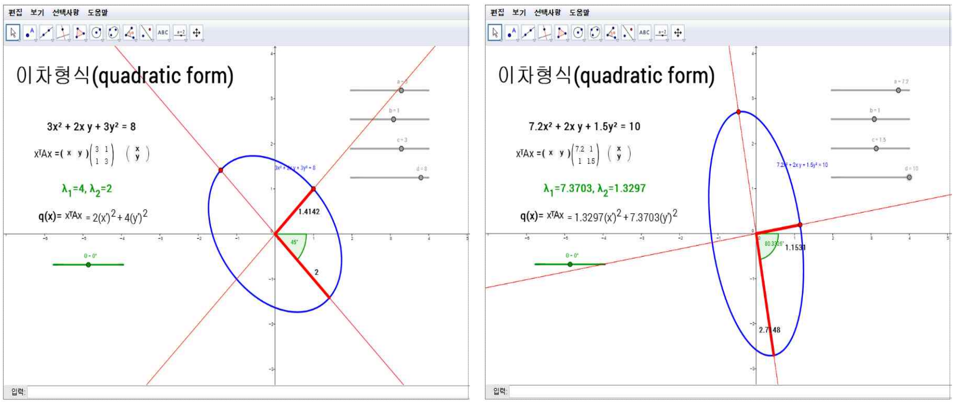


$$\frac{(x'')^2}{2} + \frac{(y'')^2}{1} = 1$$

where the $x''y''$-coordinate system is obtained by horizontally translating the $x'y'$-coordinate system 1 unit along the x'-axis. ■

[Remark] Simulation for quadratic forms

- http://www.geogebratube.org/student/m121534



Surface in 3-dimensional space

- Let

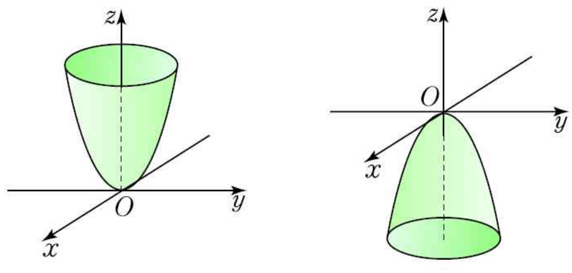
$$z = ax^2 + 2bxy + cy^2 \tag{13}$$

Then, after diagonalization, we get

$$z = \lambda_1 (x')^2 + \lambda_2 (y')^2 \tag{14}$$

in the rotated $x'y'z$-coordinate system. This enables us to identify the graph of the equation (13) in $\mathbb{R}^3$.

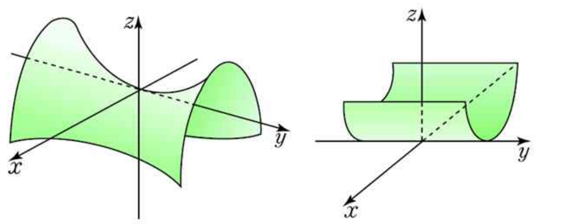
In equation (14), if both λ_1, λ_2 are positive, then the graph of equation (14) is a paraboloid opening upward (see figure (a) below). If both λ_1 and λ_2 are negative, then the graph is a paraboloid opening downward (see figure (b) below). Since the horizontal cross-section of each paraboloid is an ellipse, the above graphs are called elliptic paraboloids.



(a) (b)

Elliptic Paraboids

- In (14) if both of λ_1 and λ_2 are nonzero but have different signs, then the graphs looks like a saddle in (a) and is called a hyperbolic paraboloid. If one of λ_1 and λ_2 is zero, then the graph is parabolic cylinder in (b).



(a) hyperbolic paraboloid (b) parabolic cylinder

Example 7 Show that the graph of the following equation is an elliptic paraboloid and sketch its cross-section at $z = 50$.

$$z = 34x^2 - 24xy + 41y^2 \qquad (15)$$

Solution

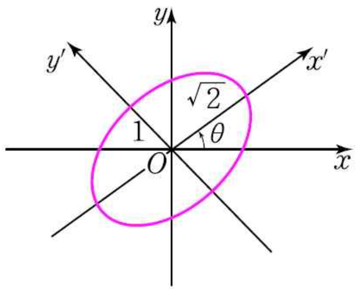
The quadratic form in (15) can be written as $z = [x\ y]\begin{bmatrix} 34 & -12 \\ -12 & 41 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$. We first find an orthogonal matrix P diagonalizing the symmetric matrix $\begin{bmatrix} 34 & -12 \\ -12 & 41 \end{bmatrix}$. It can be shown that $P = \frac{1}{5}\begin{bmatrix} 4 & -3 \\ 3 & 4 \end{bmatrix}$, and hence using $\mathbf{x} = P\mathbf{x}'$, we can transform (15) into the following:

$$z = 25(x')^2 + 50(y')^2 \qquad (16)$$

The equation (16) represents an elliptic paraboloid in the $x'y'z$-coordinate system. Note that the $x'y'$-coordinate system is obtained by rotating the xy-coordinate by angle θ counterclockwise. Hence, in $\mathbf{x}= P\mathbf{x}'$, P is given by

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

and $\theta = \tan^{-1}\left(\frac{3}{4}\right)$. Now we sketch the cross-section of equation (15) at $z = 50$. By substituting $z = 50$ into (16), we get $\frac{(x')^2}{2} + \frac{(y')^2}{1} = 1$ and hence the graph looks like the following:



□

Sage Let use Sage to graph equation (15)
http://sage.skku.edu

① Computing eigenvalues of A

```
A=matrix(2, 2, [34, -12, -12, 41])
print A.eigenvalues()
```

```
[50, 25]
```

② Computing eigenvectors of A

```
print A.eigenvectors_right()
```

```
[(50, [(1, -4/3)], 1),
(25, [(1, 3/4)], 1)]
```

③ Computing P diagonalizing A

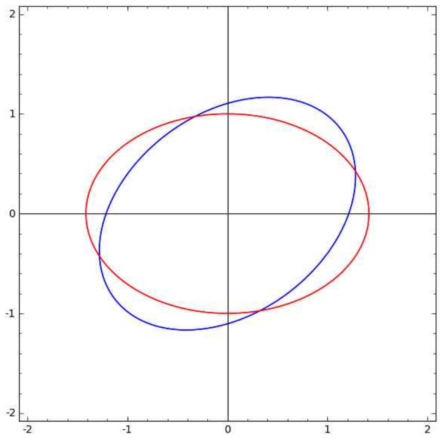
```
G=matrix([[1, 3/4], [1, -4/3]])    # Constructing a matrix whose columns
```

```
                                        # are eigenvectors
P=matrix([1/G.row(j).norm()*G.row(j) for j in range(0,2)])
# Normalizing the row vectors (The orthogonality follows from the fact #
that the eigenvalues are distinct)
P=P.transpose()  # Constructing a matrix whose columns are orthonormal
                 # eigenvectors
print P
```

```
[ 4/5  3/5]
[ 3/5 -4/5]
```

④ Sketching two ellipses simultaneously

```
var('u, v')
s=vector([u, v])
B=P.transpose()*A*P
p1=implicit_plot(s*A*s==50, (u, -2, 2), (v, -2, 2), axes='true')
p2=implicit_plot(s*B*s==50, (u, -2, 2), (v, -2, 2), color='red', axes='true')
show(p1+p2)  # Ploting two graphs simultaneously
```



■

8.5 *Applications of Quadratic forms

- Lecture Movie : http://youtu.be/cOW9qT64e0g
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-12-sec-8-5.html

By the theorem of principal axis (theorem 8.4.1), the graph of a 3D curve is shown in the form of an plane, ellipse or parabola in 2D. The specific shape is uniquely determined by signs of eigenvalues of the corresponding quadratic form. In this section, we define the sign of the quadratic form to identify the type of graph of given quadratic forms, and learn how to obtain the extrema of multivariable functions using them.

Given a system of springs and masses, there will be one quadratic form that represents the kinetic energy of the system, and another which represents the potential energy of the system in position variables. It can be found in the following websites:

- Application of Quadratic Forms and Sage:
  http://matrix.skku.ac.kr/2014-Album/Quadratic-form/
- http://ocw.mit.edu/ans7870/18/18.013a/textbook/HTML/chapter32/section09.html

8.6 SVD and generalized eigenvectors

- Lecture Movie : https://youtu.be/ejCge6Zjf1M, http://youtu.be/7-qG-A8nXmo
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-12-sec-8-6.html

We have learned that symmetric matrices are diagonalizable. We now extend the concept of diagonalization to $m \times n$ matrices (not necessarily square or symmetric) resulting in a matrix decomposition and study pseudoinverses and least squares solution using the matrix decomposition.

Theorem 8.6.1 [Singluar Value Decomposition]

Let A be an $m \times n$ real matrix. Then there exist orthogonal matrices U of order m and V of order n, and an $m \times n$ matrix Σ such that

$$U^T A V = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} = \Sigma, \tag{1}$$

where the main diagonal entries of Σ_1 are positive and listed in the monotonically decreasing order, and O is a zero-matrix. That is,

$$A = U\Sigma V^T = [\mathrm{u}_1 \mathrm{u}_2 \cdots \mathrm{u}_k \mathrm{u}_{k+1} \cdots \mathrm{u}_m] \left[\begin{array}{cccc|ccc} \sigma_1 & & & 0 & 0 & \cdots & 0 \\ & \sigma_2 & & & 0 & \cdots & 0 \\ & & \ddots & & \vdots & & \vdots \\ 0 & & & \sigma_k & 0 & \cdots & 0 \\ \hline 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{array}\right] \begin{bmatrix} \mathrm{v}_1^T \\ \mathrm{v}_2^T \\ \vdots \\ \mathrm{v}_n^T \end{bmatrix},$$

where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k > 0$.

Definition

Equation (1) is called the singular value decomposition (SVD) of A. The main diagonal entries of the matrix Σ are called the singular values of A. In addition, the columns of U are called the **left singular vectors of A** and the columns of V are called the **right singular vectors of A.**

- The following theorem shows that matrices U and V are orthogonal matrices diagonalizing AA^T and A^TA, respectively.

Theorem 8.6.2

Let the decomposition $A = U\Sigma V^T$ be the singular value decomposition (SVD) of an $m \times n$ $(m \geq n)$ matrix A where $\sigma_1, \ldots, \sigma_r$ are positive diagonal entries of Σ. Then

(1) $V^T(A^TA)V = \mathrm{diag}\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2, 0, \ldots, 0\right)_{n \times n}$.

(2) $U^T(AA^T)U = \mathrm{diag}\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2, 0, \ldots, 0\right)_{m \times m}$.

Proof Since $A = U\Sigma V^T$, it follows that $\Sigma = U^TAV$. Hence, by considering, $\Sigma^T\Sigma$ and $\Sigma\Sigma^T$, we get

$$V^TA^TAV = \Sigma' = \mathrm{diag}\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2, 0, \ldots, 0\right)_{n \times n} \quad \text{and}$$
$$U^TAA^TU = \Sigma = \mathrm{diag}\left(\sigma_1^2, \sigma_2^2, \cdots, \sigma_r^2, 0, \ldots, 0\right)_{m \times m},$$

respectively. ■

Example 1 Find the SVD of $A = \begin{bmatrix} \sqrt{3} & 2 \\ 0 & \sqrt{3} \end{bmatrix}$.

Solution

The eigenvalues of $A^TA = \begin{bmatrix} \sqrt{3} & 0 \\ 2 & \sqrt{3} \end{bmatrix}\begin{bmatrix} \sqrt{3} & 2 \\ 0 & \sqrt{3} \end{bmatrix} = \begin{bmatrix} 3 & 2\sqrt{3} \\ 2\sqrt{3} & 7 \end{bmatrix}$ are $\lambda_1 = 9$, $\lambda_2 = 1$ and hence the singular values of A are $\sigma_1 = \sqrt{\lambda_1} = 3$, $\sigma_2 = \sqrt{\lambda_2} = 1$.

A unit eigenvector of A^TA corresponding to $\lambda_1 = 9$ is $\mathrm{v}_1 = \begin{bmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix}$, and a unit eigenvector of A^TA corresponding to $\lambda_2 = 1$ is $\mathrm{v}_2 = \begin{bmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$. We can also find unit eigenvectors of AA^T:

$$\mathrm{u}_1 = \left(\frac{1}{\sigma_1}A\mathrm{v}_1 =\right)\begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}, \quad \mathrm{u}_2 = \left(\frac{1}{\sigma_2}A\mathrm{v}_2 =\right)\begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix}.$$

Hence we get

$$U = [\mathbf{u}_1\ \mathbf{u}_2] = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}, \quad V = [\mathbf{v}_1\ \mathbf{v}_2] = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}.$$

Therefore, the SVD of A is

$$A = U\Sigma V^T \Leftrightarrow \begin{bmatrix} \sqrt{3} & 2 \\ 0 & \sqrt{3} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}. \qquad \square$$

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

① Computing the singular values of A and eigenvectors of A^TA

```
A=matrix([[sqrt(3), 2], [0, sqrt(3)]])
B=A.transpose()*A
eig=B.eigenvalues()
sv=[sqrt(i) for i in eig]              # Computing singular values
print B.eigenvectors_right()           # Computing eigenvectors of A^T A
```

```
[(9, [(1, sqrt(3))], 1), (1, [(1, -1/3*sqrt(3))], 1)]
```

② Computing V^T

```
G=matrix([[1, sqrt(3)], [1, -1/3*sqrt(3)]])
Vh=matrix([1/G.row(j).norm()*G.row(j) for j in range(0,2)])
Vh=Vh.simplify()     # Transpose of V
print Vh
```

```
[          1/2  1/2*sqrt(3)]
[1/2*sqrt(3)          -1/2]
```

③ Computing eigenvectors of AA^T

```
C=A*A.transpose()
print C.eigenvectors_right()           # Computing eigenvectors of AA^T
```

```
[(9, [(1, 1/3*sqrt(3))], 1), (1, [(1, -sqrt(3))], 1)]
```

④ Computing U

```
F=matrix([[1, 1/3*sqrt(3)], [1, -sqrt(3)]])
U=matrix([1/F.row(j).norm()*F.row(j) for j in range(0,2)])
U=U.simplify().transpose()     # U
print U
```

```
[ 1/2*sqrt(3)             1/2]
[           1/2 -1/2*sqrt(3)]
```

⑤ Computing diagonal matrix S

```
S=diagonal_matrix(sv); S
```

```
[3 0]
[0 1]
```

⑥ Verifying $A = USV^T$

```
U*S*Vh
```

```
[sqrt(3)       2]
[      0 sqrt(3)]
```

■

Equivalent statement of invertible matrix on SVD

Theorem 8.6.3

Let A be an $n \times n$ matrix. Then A is a nonsingular matrix if and only if every singular value of A is nonzero.

Proof Since $\det(AA^T) = (\det A)^2$, matrix A is nonsingular if and only if AA^T is nonsingular. Hence, if A is nonsingular, then all the eigenvalues of AA^T are nonzero. By Theorem 8.6.2, the singular values of A are the square roots of the positive eigenvalues of AA^T. Hence the singular values of A are nonzero. ■

Theorem 8.6.4

Suppose $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$ are the singular values of an $m \times n$ matrix A. Then the matrix A can be expressed as follows:

$$A = \sum_{j=1}^{r} \sigma_j \mathbf{u}_j \mathbf{v}_j^T. \tag{R}$$

The equation (R) is called a rank-one decomposition of A.

- Note that the pseudoinverse of a matrix is important in the study of the least squares solutions for optimization problems.

- We can express an $n \times n$ nonsingular matrix A using the SVD

$$A = U\Sigma V^T. \tag{2}$$

Note that all of U, Σ, V are $n \times n$ nonsingular matrices and , in particular, U, V are orthogonal matrices. Hence the inverse of A can be expressed as

$$A^{-1} = V\Sigma^{-1}U^T. \tag{3}$$

- If A is not a square matrix or A is singular, then (3) does not give an inverse of A. However, we can construct a pseudoinverse $A^\dagger$ of A by putting Σ in (2) into the form $\Sigma = \begin{bmatrix} \Sigma_1 & O \\ O & O \end{bmatrix}$(where Σ_1 is nonsingular).

Definition [Pseudo-Inverse]

For an $m \times n$ matrix A the $n \times m$ matrix $A^\dagger = V\Sigma' U^T$ is called a pseudo-inverse of A, where U, V are orthogonal matrices in the SVD of A and Σ' is

$$\Sigma' = \begin{bmatrix} \Sigma_1^{-1} & O \\ O & O \end{bmatrix} \text{ (where } \Sigma_1 \text{ is nonsingular).}$$

- We read $A^\dagger$ as A 'dagger.' If $A = O$, then we define $A^\dagger = O$.

- Truncated SVD

http://langvillea.people.cofc.edu/DISSECTION-LAB/Emmie'sLSI-SVDModule/p5module.html

What is a truncated SVD?

We learned that singular value decomposition factors any matrix A so that $A = U\,S\,V^T$. Let's take a closer look at the matrix S. Remember $S = \begin{bmatrix} \Sigma_1 & O \\ O & O \end{bmatrix}$ is a diagonal matrix where $\Sigma_1 = \begin{bmatrix} \sigma_1 & & 0 & \\ & \sigma_2 & & \\ \vdots & \vdots & \vdots & \\ 0 & & \cdots\ \sigma_r & \end{bmatrix}$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are the singular values of the matrix A. A full rank decomposition of A is usually denoted by $A_r = U_r\ \Sigma_1\ V_r^T$ where U_r and V_r are the matrices obtained by taking the first r columns of U and V, respectively. We can find a k-rank approximation (or truncated SVD) to A by taking only the first k largest singular values and the first k columns of U and V.

Example 2

Find a pseudo-inverse of $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Solution

We first compute the (truncated) SVD[1] of A:
http://matrix.skku.ac.kr/2014-Album/MC.html

$$A = [\mathrm{u}_1\ \mathrm{u}_2]\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}\begin{bmatrix} \mathrm{v}_1^T \\ \mathrm{v}_2^T \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{6}}{3} & 0 \\ \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} \end{bmatrix}\begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}.$$

Then

$$A^{\dagger} = [\mathrm{v}_1\ \mathrm{v}_2]\begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ 0 & \frac{1}{\sigma_2} \end{bmatrix}\begin{bmatrix} \mathrm{u}_1^T \\ \mathrm{u}_2^T \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \frac{\sqrt{6}}{3} & \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} \\ 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

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

■

1) http://langvillea.people.cofc.edu/DISSECTION-LAB/Emmie'sLSI-SVDModule/p5module.html

- http://matrix.skku.ac.kr/RPG_English/8-MA-pseudo-inverse.html

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

```
A=matrix(RDF,[[1,1],[0,1],[1,0]])
import numpy as np
Pinv=matrix(np.linalg.pinv(A))
Pinv
```

```
[ 0.333333333333 -0.333333333333  0.666666666667]
[ 0.333333333333  0.666666666667 -0.333333333333]
```

- If $A=[a_{ij}]_{m\times n}$ has $\text{rank}(A)=n$, then A is said to be of the full column rank. If A has full column rank, then A^TA is nonsingular. If A is nonsingular, then the pseudo-inverse of A is equal to A^{-1}.

Theorem 8.6.5

If an $m\times n$ matrix A has full column rank, then the pseudo-inverse of A is

$$A^{\dagger}=(A^TA)^{-1}A^T.$$

Proof Let $A=U\Sigma V^T$ be the SVD of A. Then $\Sigma=\begin{bmatrix}\Sigma_1\\ O\end{bmatrix}$ where Σ_1 is nonsingular. Then

$$A^TA=(V\Sigma^TU^T)(U\Sigma V^T)=V\Sigma_1^2V^T.$$

Since A has full column rank, A^TA is nonsingular and matrix V is an $n\times n$ orthogonal matrix. Hence $(A^TA)^{-1}=V\Sigma_1^{-2}V^T$ and

$$\begin{aligned}(A^TA)^{-1}A^T&=(V\Sigma_1^{-2}V^T)(U\Sigma V^T)^T=(V\Sigma_1^{-2}V^T)(V\Sigma^TU^T)\\&=(V\Sigma_1^{-2}V^T)(V[\Sigma_1\ \ O]\,U^T)=V[\Sigma_1^{-1}\ \ O]\,U^T=A^{\dagger}.\end{aligned}$$

■

Example 3 Find the pseudo-inverse of A using theorem 8.6.5.

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

Solution

Since A has full column rank,

$$A^TA=\begin{bmatrix}1&0&1\\1&1&0\end{bmatrix}\begin{bmatrix}1&1\\0&1\\1&0\end{bmatrix}=\begin{bmatrix}2&1\\1&2\end{bmatrix}$$ is nonsingular and

$$A^{\dagger}=(A^TA)^{-1}A^T=\begin{bmatrix}\frac{2}{3}&-\frac{1}{3}\\-\frac{1}{3}&\frac{2}{3}\end{bmatrix}\begin{bmatrix}1&0&1\\1&1&0\end{bmatrix}=\begin{bmatrix}\frac{1}{3}&-\frac{1}{3}&\frac{2}{3}\\\frac{1}{3}&\frac{2}{3}&-\frac{1}{3}\end{bmatrix}.$$ ■

Theorem 8.6.6

If $A^{\dagger}$ is a pseudo-inverse of A, then the following hold:

(1) $AA^{\dagger}A=A$

(2) $A^{\dagger}AA^{\dagger}=A^{\dagger}$

(3) $(AA^{\dagger})^T=AA^{\dagger}$

(4) $(A^{\dagger}A)^T=A^{\dagger}A$

(5) $(A^T)^{\dagger}=(A^{\dagger})^T$

(6) $A^{\dagger\dagger}=A$.

Proof (Exercise)

[Remark]

A pseudo-inverse provides a tool for solving a least sqaures problem. It is known that the least squares solution to the linear system $A\mathbf{x}=\mathbf{b}$ is the solution to the normal equation $A^TA\mathbf{x}=A^T\mathbf{b}$. If A has full column rank, then the matrix A^TA is nonsingular and hence

$$\mathbf{x}=(A^TA)^{-1}A^T\mathbf{b}=A^{\dagger}\mathbf{b}.$$

This means that if A has full column rank, the least squares solution to $A\mathbf{x}=\mathbf{b}$ is the pseudo-inverse $A^{\dagger}$ times the vector $\mathbf{b}$.

Theorem 8.6.7

Let A be an $m \times n$ matrix and $\mathbf{b}$ be a vector in $\mathbb{R}^n$. Then $\mathbf{x} = A^\dagger \mathbf{b}$ is the least squares solution to $A\mathbf{x} = \mathbf{b}$.

Proof Let $A = U\Sigma V^T$ be the SVD of A with $\Sigma = \begin{bmatrix} \Sigma_1 & O \\ O & O \end{bmatrix}$ (Σ_1 is nonsingular). Then $A^\dagger = V\Sigma' U^T = V\begin{bmatrix} \Sigma_1^{-1} & O \\ O & O \end{bmatrix} U^T$ and hence $A^\dagger \mathbf{b} = V\Sigma' U^T \mathbf{b}$, Since

$$\begin{aligned}(A^T A)A^\dagger \mathbf{b} &= V\Sigma^T \Sigma V^T V \Sigma' U^T \mathbf{b} = V\Sigma^T \Sigma \Sigma' U^T \mathbf{b} \\ &= V\begin{bmatrix} \Sigma_1^2 & O \\ O & O \end{bmatrix}\begin{bmatrix} \Sigma_1^{-1} & O \\ O & O \end{bmatrix} U^T \mathbf{b} = V\Sigma^T U^T \mathbf{b} = A^T \mathbf{b},\end{aligned}$$

it follows that $\mathbf{x} = A^\dagger \mathbf{b}$ satisfies $A^T A\mathbf{x} = A^T \mathbf{b}$. ■

Example 4 Find the least squares line passing through the four points $(0,1), (1,3), (2,4), (3,4)$.

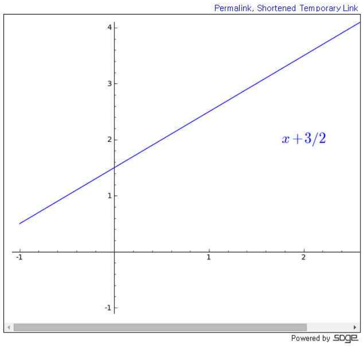
Solution

Let $y = mx + b$ be an equation for the line that fits to the points $(0,1), (1,3), (2,4), (3,4)$. Then, by letting $\mathbf{x} = (b \quad m)^{\mathrm{T}}$, the given condition can be written as the linear system $A\mathbf{x} = \mathbf{b}$ for which

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 4 \\ 4 \end{bmatrix}.$$

Since A has full column rank, we get $A^\dagger = (A^T A)^{-1} A^T$ which is

$A^\dagger = \begin{bmatrix} \frac{7}{10} & \frac{-3}{10} \\ -\frac{3}{10} & \frac{2}{10} \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix}$. Hence $\mathbf{x} = A^\dagger \mathbf{b} = \begin{bmatrix} \frac{3}{2} \\ 1 \end{bmatrix}$. Therefore, the least squares line is given by $y = x + \dfrac{3}{2}$. ■




```
var('x, y')
p1=plot(x + 3/2, x, -1,3, color='blue');
p2 = text("$x+ 3/2 $", (2,2), fontsize=20, color='blue')
show(p1+p2, ymax=4, ymin=-1)
```
in http://matrix.skku.ac.kr/Cal-Book/part1/CS-Sec-1-3.htm

Team <3D Math> comprises Professor Sang-Gu LEE, and 3 mathematics major students including Jaeyoon LEE, Victoria LANG, Youngjun LIM, won the prize with 'DIY Math Tools with 3D printer' for the Korea Science and Technology Idea Competition 2014 co-organized by the Korea Foundation for the Advancement of Science and Creativity, the Ministry of Science, ICT and Future Planning, the National Museum of Science and YTN.
https://www.facebook.com/skkuscience

8.7 Complex eigenvalues and eigenvectors

- Lecture Movie : http://youtu.be/8_uNVj_OlAk, http://youtu.be/Ma2er-9LC_g
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-13-sec-8-7.html

We have so far focused on real eigenvalues and real eigenvectors. However, real square matrices can have complex eigenvalues and eigenvectors. In this section, we introduce complex vector spaces, complex matrices, complex eigenvalues and complex eigenvectors.

Complex vector spaces

Definition [Complex Vector Space]

The set of vectors with n complex components is denoted by

$$C^n = \{(z_1, z_2, \ldots, z_n) \mid z_k \in C,\ k = 1, 2, \ldots, n\}.$$

If we define the vector addition and the scalar multiple of a vector in C^n similar to those for $\mathbb{R}^n$, then C^n is a vector space over C and its dimension is equal to n.

- If

$$\mathbf{e}_1 = (1, 0, \ldots, 0),\ \mathbf{e}_2 = (0, 1, 0, \ldots, 0),\ \ldots,\ \mathbf{e}_n = (0, \ldots, 0, 1),$$

then a vector $\mathbf{v}$ in C^n can be expressed as $\mathbf{v} = z_1\mathbf{e}_1 + z_2\mathbf{e}_2 + \cdots + z_n\mathbf{e}_n$ where z_i's are complex numbers, and the set $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis for C^n. This basis is called the standard basis for C^n.

- For a complex number $z = a + bi$, $\bar{z} = a - bi$ is called the conjugate of z and $|z| = \sqrt{a^2 + b^2}$ is called the modulus of z. Furthermore, if we denote a complex number z as $z = r(\cos\theta + i\sin\theta)$, then $r = |z|$ and $\tan\theta = \frac{b}{a}$. For a complex vector $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, we define its conjugate as $\bar{\mathbf{u}} = (\overline{u_1}, \overline{u_2}, \ldots, \overline{u_n})$.

- [Example] http://matrix.skku.ac.kr/RPG_English/9-VT-conjugate.html

Inner product

Definition [Inner Product]

Let $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ be vectors in C^n. Then

$$\mathbf{u} \cdot \mathbf{v} = \overline{v_1} u_1 + \overline{v_2} u_2 + \cdots + \overline{v_n} u_n = <\mathbf{u}, \mathbf{v}>$$

satisfies the following properties:

(1) $<\mathbf{u}, \mathbf{v}> = \overline{<\mathbf{v}, \mathbf{u}>}$
(2) $<\mathbf{u}+\mathbf{v}, \mathbf{w}> = <\mathbf{u}, \mathbf{w}> + <\mathbf{v}, \mathbf{w}>$
(3) $<c\mathbf{u}, \mathbf{v}> = c<\mathbf{u}, \mathbf{v}>$
(4) $<\mathbf{v}, \mathbf{v}> \geq 0$, in particular, $<\mathbf{v}, \mathbf{v}> = 0 \Leftrightarrow \mathbf{v} = 0$

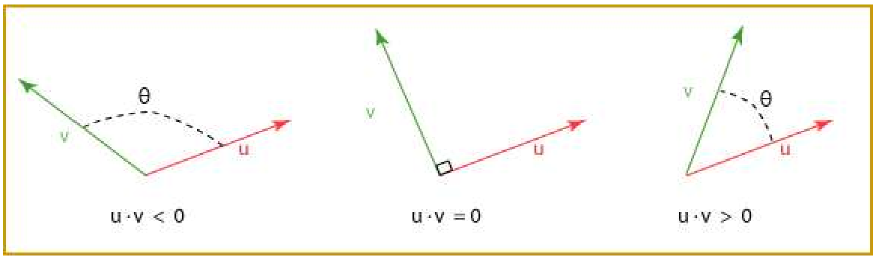
The inner product $\mathbf{u} \cdot \mathbf{v}$ is called the Euclidean inner product for the vector space C^n.

Definition

Let $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ be vectors in C^n. Then, using the Euclidean inner product $\mathbf{u} \cdot \mathbf{v}$, we can define the Euclidean norm $\|\mathbf{u}\|$ of $\mathbf{u}$ and the Euclidean distance $d(\mathbf{u}, \mathbf{v})$ between $\mathbf{u}$ and $\mathbf{v}$ as the following:

(1) $\|\mathbf{u}\| = (\mathbf{u} \cdot \mathbf{u})^{\frac{1}{2}} = \sqrt{|u_1|^2 + |u_2|^2 + \cdots + |u_n|^2}$.
(2) $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{|u_1 - v_1|^2 + |u_2 - v_2|^2 + \cdots + |u_n - v_n|^2}$.

- If $\mathbf{u} \cdot \mathbf{v} = 0$, then we say that $\mathbf{u}$ and $\mathbf{v}$ are orthogonal to each other.



Example 1 For vectors $\mathbf{u}=(2i,\ 0,\ 1+3i)$, $\mathbf{v}=(2-i,\ 0,\ 1+3i)$, compute the Euclidean inner product and their Euclidean distance.

Solution

$$\begin{aligned}\mathbf{u}\cdot\mathbf{v} &= \overline{(2-i)}\,(2i) + 0\cdot 0 + \overline{(1+3i)}\,(1+3i)\\ &= (2i)(2+i)+0+(1+3i)(1-3i) = 4i+2i^2+1-9i^2 = 8+4i\end{aligned}$$

$$\begin{aligned}d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\| &= \sqrt{|2i-(2-i)|^2+|0-0|^2+|(1+3i)-(1+3i)|^2}\\ &= \sqrt{|-2+3i|^2+0+0} = \sqrt{4+9} = \sqrt{13}\end{aligned}$$

□

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

```
u=vector([2*I, 0, 1+3*I])     # I is the imaginary unit.
v=vector([2-I, 0, 1+3*I])
print v.hermitian_inner_product(u)   # Taking the conjugate for v
# < u, v > = v.hermitian_inner_product(u)
print (u-v).norm()
```

```
4*I + 8
sqrt(13)
```

■

Complex Eigenvalues and Eigenvectors of Real Matrices

We should first define complex eigenvalues and complex eigenvectors along with example.

Theorem 8.7.1

If λ is a complex eigenvalue of an $n\times n$ real matrix A and $\mathbf{x}$ is its corresponding eigenvector of A, then the complex conjugate $\overline{\lambda}$ of λ is also an eigenvalue of A and $\overline{\mathbf{x}}$ is an eigenvector corresponding to $\overline{\lambda}$.

Proof Since an eigenvector is a nonzero vector, $\mathbf{x}\neq\mathbf{0}$ and $\overline{\mathbf{x}}\neq\mathbf{0}$. Since $A\mathbf{x}=\lambda\mathbf{x}$ and A is real (i.e., $\overline{A}=A$), it follows that $A\overline{\mathbf{x}}=\overline{A\mathbf{x}}=\overline{\lambda\mathbf{x}}=\overline{\lambda}\,\overline{\mathbf{x}}$. ■

Eigenvalues of Real Symmetric Matrices

Theorem 8.7.2

If A is a real symmetric matrix, then all the eigenvalues of A are real numbers.

Proof Let λ be an eigenvalue of A, that is, there exists a nonzero vector $\mathbf{x}$ such that $A\mathbf{x}=\lambda\mathbf{x}$. By multiplying both sides by $\mathbf{x}^* = \overline{\mathbf{x}}^T$ on the left-hand side, we get $\mathbf{x}^*A\mathbf{x}=\mathbf{x}^*(\lambda\mathbf{x})=\lambda\mathbf{x}^*\mathbf{x}=\lambda(\mathbf{x}\cdot\ \mathbf{x})=\lambda\|\mathbf{x}\|^2$. Hence $\lambda=\dfrac{\mathbf{x}^*A\mathbf{x}}{\|\mathbf{x}\|^2}$. Since $\|\mathbf{x}\|^2$ is a nonzero real number, we just need to show that $\mathbf{x}^*A\mathbf{x}$ is a real number. Note that

$$\overline{\mathbf{x}^*A\mathbf{x}}=\overline{\mathbf{x}^*}\,\overline{A\mathbf{x}}=\mathbf{x}^T(\overline{A\mathbf{x}})=(\overline{A\mathbf{x}})^T\mathbf{x}=(\overline{A}\,\overline{\mathbf{x}})^T\mathbf{x}=\mathbf{x}^*A^*\mathbf{x}=\mathbf{x}^*A\mathbf{x}.$$

Therefore, $\mathbf{x}^*A\mathbf{x}$ is a real number. ■

Example 2 Show that the eigenvalues of $C=\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ are $\lambda=a\pm bi$. In addition show that if $(a,b)\neq(0,0)$, then C can be decomposed into

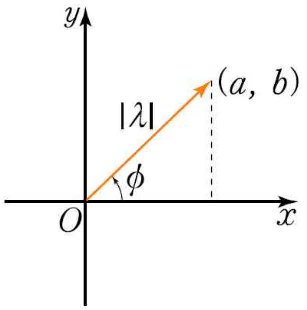
$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}=\begin{bmatrix} |\lambda| & 0 \\ 0 & |\lambda| \end{bmatrix}\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix},$$

where ϕ is the angle between the x-axis and the line passing through the origin and the point $(a,\,b)$.

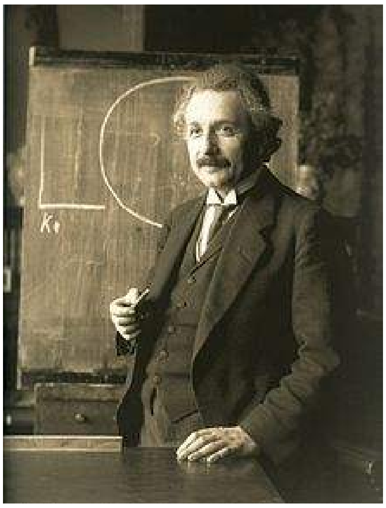
Solution

Since the characteristic equation of C is $(\lambda-a)^2+b^2=0$, the eigenvalues of C are $\lambda=a\pm bi$. If $(a,b)\neq(0,0)$, then $a=|\lambda|\cos\phi$, $b=|\lambda|\sin\phi$. Therefore,

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}=\begin{bmatrix} |\lambda| & 0 \\ 0 & |\lambda| \end{bmatrix}\begin{bmatrix} \dfrac{a}{|\lambda|} & -\dfrac{b}{|\lambda|} \\ \dfrac{b}{|\lambda|} & \dfrac{a}{|\lambda|} \end{bmatrix}=\begin{bmatrix} |\lambda| & 0 \\ 0 & |\lambda| \end{bmatrix}\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}.$$ ■



“Pure mathematics is, in its way, the poetry of logical ideas."



Albert Einstein (1879-1955)
http://en.wikipedia.org/wiki/Albert_Einstein
He developed the general theory of relativity, one of the two pillars of modern physics (alongside quantum mechanics)

8.8 Hermitian, Unitary, Normal Matrices

- Lecture Movie : http://youtu.be/8_uNVj_OlAk, http://youtu.be/GLGwj6tzd60
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-13-sec-8-8.html

We used M_n to denote the set of all $n\times n$ real matrices. In this section, we introduce $M_n(C)$ to denote the set of all $n\times n$ complex matrices. Symmetric matrices and orthogonal matrices in M_n can be generalized to be Hermitian matrices and unitary matrices in $M_n(C)$. We shall further study the diagonalization of Hermitian and Unitary matrices.

Conjugate Transpose

Definition [Conjugate Transpose]

For a matrix $A=[a_{ij}]\in M_{m\times n}(C)$, $\overline{A}$ is defined by

$$\overline{A} = [\overline{a_{ij}}]\in M_{m\times n}(C).$$

The transpose $\overline{A}^T$ of the complex conjugate of A is called the conjugate transpose and is denoted by A^*, that is, $A^*=\overline{A}^T=[\overline{a_{ji}}]_{n\times m}$.

[Remark]

- The Euclidean inner product in C^n: $\mathbf{u}\cdot\mathbf{v}=\mathbf{v}^*\mathbf{u}$, $\|\mathbf{u}\|^2=\mathbf{u}^*\mathbf{u}$
- If a matrix A is real, then $A^*=A^T$.

Example 1

For matrices $A=\begin{bmatrix}1+i & -i & 0\\ 2 & 3-2i & i\end{bmatrix}$, $B=\begin{bmatrix}1 & 1+2i\\ 1-2i & 0\end{bmatrix}$, $C=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}$, their conjugate transposes are

$$A^* = \begin{bmatrix} 1-i & 2 \\ i & 3+2i \\ 0 & -i \end{bmatrix}, \; B^* = \begin{bmatrix} 1 & 1+2i \\ 1-2i & 0 \end{bmatrix}, \; C^* = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}.$$ ■

Theorem 8.8.1 [Properties of Conjugate Transpose]

For complex matrices A, B and a complex number c, the following hold:

(1) $(A^*)^* = A$.
(2) $(A+B)^* = A^* + B^*$.
(3) $(cA)^* = \bar{c} A^*$.
(4) $(AB)^* = B^* A^*$.

Proof of the above theorem is easy to verify and left as exercises.

Hermitian Matrix

Definition [Hermitian Matrix]

If a complex square matrix A satisfying $A = A^*$, A is called a **Hermitian matrix.**

Example 2 In Example 1, $A \neq A^*$ and hence A is not Hermitian. However, since $B = B^*$, B is Hermitian. ■

Theorem 8.8.2 [Properties of Hermitian Matrix]

Suppose $A \in M_n(C)$ is Hermitian. Then the following hold:

(1) For any vector $\mathbf{x} \in C^n$, the product $\mathbf{x}^* A \mathbf{x}$ is a real number.
(2) Every eigenvalue of A is a real number.
(3) Eigenvectors of A corresponding to distinct eigenvalues are orthogonal to each other.

http://people.math.gatech.edu/~meyer/MA6701/module11.pdf

Example 3 Let $A=\begin{bmatrix}1 & 0 & 0\\ 0 & 2 & 1-i\\ 0 & 1+i & 0\end{bmatrix}$. Since $A=A^*$, A is Hermitian. The characteristic equation of A is $|A-\lambda I|=(\lambda-1)(\lambda^2-2\lambda-2)=0$ and hence the eigenvalues of A are $\lambda=1,\ 1+\sqrt{3},\ 1-\sqrt{3}$, which confirms that all the eigenvalues of a Hermitian matrix A are real numbers. Furthermore, it can be shown that the eigenvectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$

$$\mathbf{x}=(1,\,0,\,0),\ \ \mathbf{y}=\left(0,\ \left(-\frac{1}{2}+\frac{i}{2}\right)(-1+\sqrt{3}),\ 1\right),\ \ \mathbf{z}=\left(0,\ \left(\frac{1}{2}-\frac{i}{2}\right)(1+\sqrt{3}),\ 1\right)$$

corresponding to $\lambda=1$, $\lambda=1+\sqrt{3}$, and $\lambda=1-\sqrt{3}$, respectively, are orthogonal to each other. ■

Skew-Hermitian Matrices

Definition **[Skew-Hermitian Matrix]**

If a complex square matrix A satisfies $A=-A^*$, then A is called a **skew-Hermitian matrix.**

Example 4 It can be verified that both matrices A and B below are skew-Hermitian:

$$A=\begin{bmatrix}-i & -5i\\ -5i & 3i\end{bmatrix},\ \ B=\begin{bmatrix}0 & 0 & i\\ 0 & i & 0\\ i & 0 & 0\end{bmatrix},\ \ A^*=\begin{bmatrix}i & 5i\\ 5i & -3i\end{bmatrix}=-A \text{ and } B^*=-B$$ ■

- Each matrix $A\in M_n(C)$ can be expressed as $A=H+K$, where H is Hermitian and K is skew-Hermitian. In particular, since $A+A^*$ is Hermitian and $A-A^*$ is skew-Hermitian, every complex square matrix A can be rewritten as

$$A=\frac{1}{2}(A+A^*)+\frac{1}{2}(A-A^*).$$

Unitary Matrices

Definition **[Unitary Matrix]**

If matrix $U \in M_n(C)$ satisfies $U^* U = I_n$, then U is called a **unitary matrix.** If U is unitary, then $U^* = U^{-1}$. In addition, if the jth column vector of U is denoted by $\mathbf{u}_j$, then

$$\mathbf{u}_i \cdot \mathbf{u}_j = <\mathbf{u}_i, \mathbf{u}_j> = \mathbf{u}_j^* \mathbf{u}_i = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases}.$$

Therefore, U is a unitary matrix if and only if the columns of U form an orthonormal set in C^n.

Example 5 Show that the following matrix A is unitary:

$$A = \frac{1}{2}\begin{bmatrix} 1+i & 1+i \\ 1-i & -1+i \end{bmatrix}$$

Solution

Since $A^* = \frac{1}{2}\begin{bmatrix} 1-i & 1+i \\ 1-i & -1-i \end{bmatrix}$, the product $A^*A = \frac{1}{4}\begin{bmatrix} 1-i & 1+i \\ 1-i & -1-i \end{bmatrix}\begin{bmatrix} 1+i & 1+i \\ 1-i & -1+i \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$. Hence $A = [\mathbf{a}_1 : \mathbf{a}_2]$ is a unitary matrix. We can also show that

$$\mathbf{a}_i \cdot \mathbf{a}_j = \mathbf{a}_j^* \mathbf{a}_i = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases}.$$

For example, $\mathbf{a}_1 \cdot \mathbf{a}_2 = \mathbf{a}_2^* \mathbf{a}_1 = \frac{1}{4}[1+i \ \ -1+i]\begin{bmatrix} 1+i \\ 1-i \end{bmatrix} = 0$. ■

Theorem 8.8.3 [Properties of a Unitary Matrix]

Suppose C^n has the Euclidean inner product and U is a unitary matrix. Then the following hold:

(1) For $\mathbf{x}, \mathbf{y} \in C^n$, $(U\mathbf{x})\cdot (U\mathbf{y}) = (\mathbf{x}\cdot \mathbf{y})$, which implies $\|U\mathbf{x}\| = \|\mathbf{x}\|$.
(2) If λ is an eigenvalue of U, then $|\lambda| = 1$.
(3) Eigenvectors of U corresponding to distinct eigenvalues are orthogonal to each other.

- The property $\|U\mathbf{x}\| = \|\mathbf{x}\|$ of a unitary matrix U shows that a unitary matrix is an isometry, preserving the norm.

Unitary Similarity and Unitarily Diagonalizable Matrices

Definition [Unitary Similarity and Unitary Diagonalization]

For matrices $A, B \in M_n(C)$, if there exists a unitary matrix U such that $U^* A U = B$, then we say that A and B are unitarily similar to each other. Furthermore, if $A \in M_n(C)$ is unitarily similar to a diagonal matrix, then A is called unitarily diagonalizable.

Example 6 Let $A = \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix}$ and $U = \frac{1}{\sqrt{2}}\begin{bmatrix} i & -1 \\ 1 & -i \end{bmatrix}$. Then it can be checked that U is a unitary matrix and $U^* A U = \frac{1}{2}\begin{bmatrix} -i & 1 \\ -1 & i \end{bmatrix}\begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix}\begin{bmatrix} i & -1 \\ 1 & -i \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$.
Therefore, A is unitarily diagonalizable. ■

- If $A \in M_n(C)$ is unitarily diagonalizable, then there exists a unitary matrix U such that $U^* A U = D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and hence $AU = UD$. Letting $U = [U^{(1)} : U^{(2)} : \cdots : U^{(n)}]$, we get,

$$[AU^{(1)} : AU^{(2)} : \cdots : AU^{(n)}] = AU = UD = [\lambda_1 U^{(1)} : \lambda_2 U^{(2)} : \cdots : \lambda_n U^{(n)}].$$

- This implies that the column $U^{(i)}$ of the unitary matrix U is a unit eigenvector of A corresponding to the eigenvalue λ_i.

Example 7

Find a unitary matrix U diagonalizing matrix $A = \begin{bmatrix} 2 & 1-i \\ 1+i & 1 \end{bmatrix}$.

Solution

The eigenvalues of A are $\lambda_1 = 0, \lambda_2 = 3$ and their corresponding eigenvectors are

$$\lambda_1 = 0 \ \Rightarrow \ \mathrm{x}_1 = \begin{bmatrix} -1 \\ 1+i \end{bmatrix}, \quad \lambda_2 = 3 \ \Rightarrow \ \mathrm{x}_2 = \begin{bmatrix} 1-i \\ 1 \end{bmatrix}.$$

Letting $\mathrm{u}_1 = \dfrac{\mathrm{x}_1}{\|\mathrm{x}_1\|} = \dfrac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1+i \end{bmatrix}$, $\mathrm{u}_2 = \dfrac{\mathrm{x}_2}{\|\mathrm{x}_2\|} = \dfrac{1}{\sqrt{3}} \begin{bmatrix} 1-i \\ 1 \end{bmatrix}$

and $U = [\mathrm{u}_1 \, \mathrm{u}_2] = \dfrac{1}{\sqrt{3}} \begin{bmatrix} -1 & 1-i \\ 1+i & 1 \end{bmatrix}$, it follows that

$$U^* A U = \frac{1}{3} \begin{bmatrix} -1 & 1-i \\ 1+i & 1 \end{bmatrix} \begin{bmatrix} 2 & 1-i \\ 1+i & 1 \end{bmatrix} \begin{bmatrix} -1 & 1-i \\ 1+i & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix},$$

where U is a unitary matrix. ■

Schur's Theorem

- Transforming a complex square matrix into an upper triangular matrix

Theorem 8.8.4 [Schur's Theorem]

A square matrix A is unitarily similar to an upper triangular matrix whose main diagonal entries are the eigenvalues of A. That is, there exists a unitary matrix U and an upper triangular matrix T such that

$$U^* A U = T = [t_{ij}] \in M_n(C), \ t_{ij} = 0 (i > j),$$

where t_{ii}'s are eigenvalues of A.

Proof Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of A. We prove this by mathematical induction. First, if $n = 1$, then the statement holds because $A = [\lambda_1]$. We now assume that the statement is true for any square matrix of order less than or equal to $n-1$.

① Let x_1 be an eigenvector corresponding to eigenvalue λ_1.

② By the Gram-Schmidt Orthonormalization, there exists an orthonormal basis for C^n including x_1, say $S=\{\mathrm{x}_1, \mathrm{z}_2, \ldots, \mathrm{z}_n\}$.

③ Since S is orthonormal, the matrix $U_0 \equiv [\mathrm{x}_1 \ : \mathrm{z}_2 : \ \cdots \ : \mathrm{z}_n]$ is a unitary matrix. In addition, since $A\mathrm{x}_1 = \lambda_1 \mathrm{x}_1$, the first column of AU_0 is $\lambda_1 \mathrm{x}_1$. Hence $U_0^*(AU_0)$ is of the following form:

$$U_0^* A U_0 = \left[\begin{array}{c:c} \lambda_1 & * \\ \hdashline O & A_1 \end{array}\right],$$

where $A_1 \in M_{n-1}(C)$. Since $|\lambda I_n - A| = (\lambda - \lambda_1)\,|\lambda I_{n-1} - A_1|$, the eigenvalues of A_1 are $\lambda_2, \lambda_3, \cdots, \lambda_n$.

④ By the induction hypothesis, there exists a unitary matrix $\widehat{U}_1 \in M_{n-1}(C)$ such that

$$\widehat{U}_1^* A_1 \widehat{U}_1 = \begin{bmatrix} \lambda_2 & & * \\ 0 & & \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}.$$

⑤ Letting $U_1 \equiv \left[\begin{array}{c:c} 1 & 0 \ \cdots \ 0 \\ \hdashline 0 & \\ \vdots & \widehat{U}_1 \\ 0 & \end{array}\right] \in M_n(C)$, we get

$$(U_0U_1)^* A (U_0U_1) = U_1^* U_0^* A U_0 U_1 = \begin{bmatrix} \lambda_1 & & & * \\ & & & \\ 0 & \lambda_2 & & * \\ 0 & 0 & & \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Since $U \equiv U_0U_1$ is a unitary matrix, the result follows. ■

● [Lecture on this proof] http://youtu.be/lL0VdTStJDM

● Not every square matrix is unitarily diagonalizable. (see Chapter 10)

[Issai Schur(1875-1941, Germany)] http://en.wikipedia.org/wiki/Issai_Schur

Normal matrix

Definition [Normal Matrix]

If matrix $A \in M_n(C)$ satisfies

$$AA^* = A^*A,$$

then A is called **normal matrix**.

Example 8 It can be shown that the following matrices A and B are normal:

$$A = \begin{bmatrix} \frac{-1+i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{1+i}{2} \end{bmatrix} \quad B = \begin{bmatrix} 2+2i & i & 1-i \\ i & -2i & 1-3i \\ 1-i & 1-3i & -3+8i \end{bmatrix}$$

■

Example 9 A Hermitian matrix A satisfies $A = A^*$ and hence $AA^* = AA = A^*A$. This implies that any Hermitian matrix is normal. In addition, since a unitary matrix B satisfies $BB^* = I_n = B^*B$, it is a normal matrix. ■

Equivalent Conditions for a Matrix to be Normal

Theorem 8.8.5

For matrix $A \in M_n(C)$, the following are equivalent:

(1) A is unitarily diagonalizable.
(2) A is a normal matrix.
(3) A has n orthonormal eigenvectors.

Example 10

Let $A = \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix}$ and $U = \frac{1}{\sqrt{2}} \begin{bmatrix} i & -1 \\ 1 & -i \end{bmatrix}$.

Show that A is a normal matrix and the columns of U are orthonormal eigenvectors of A.

Solution

Since $A = \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix} = A^*$, $AA^* = A^*A$ and hence A is a normal matrix.

Letting $U = \begin{bmatrix} \frac{i}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \end{bmatrix} \equiv [\mathbf{u}_1 : \mathbf{u}_2]$, we get

$$A\mathbf{u}_1 = \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix} \begin{bmatrix} \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = 3 \begin{bmatrix} \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = 3\mathbf{u}_1,$$

$$A\mathbf{u}_2 = \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} \end{bmatrix} = 1\mathbf{u}_2.$$

Thus $\mathbf{u}_1$ and $\mathbf{u}_2$ are eigenvectors of A. In addition, since $\|\mathbf{u}_1\| = \|\mathbf{u}_2\| = 1$ and $\mathbf{u}_1 \cdot \mathbf{u}_2 = \mathbf{u}_2^* \mathbf{u}_1 = 0$, $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthonormal eigenvectors of A. ■

Example 11

Unitarily diagonalize $A = \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix}$.

Solution

Note that matrix A is Hermitian and its eigenvalues are $\lambda_1 = 3$, $\lambda_2 = 1$.

An eigenvector corresponding to $\lambda_1 = 3$ is $\mathbf{x}_1 = \begin{bmatrix} i \\ 1 \end{bmatrix}$. By normalizing it, we get $\mathbf{u}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} i \\ 1 \end{bmatrix}$. Similarly, we can get a unit eigenvector $\mathbf{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ -\mathrm{i} \end{bmatrix}$ corresponding to $\lambda_2 = 1$.

Taking $U = [\mathbf{u}_1 : \mathbf{u}_2] = \frac{1}{\sqrt{2}} \begin{bmatrix} i & -1 \\ 1 & -i \end{bmatrix}$, we get $U^*AU = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$. ■

[Remark]

Although not every matrix A is diagonalizable, using the Schur's Theorem, we can obtain an upper triangular matrix J_A (close to a diagonal matrix) similar to A. The upper triangular matrix J_A is called the Jordan canonical form of A. The Jordan canonical form will be discussed in Chapter 10.

8.9 *Linear system of differential equations

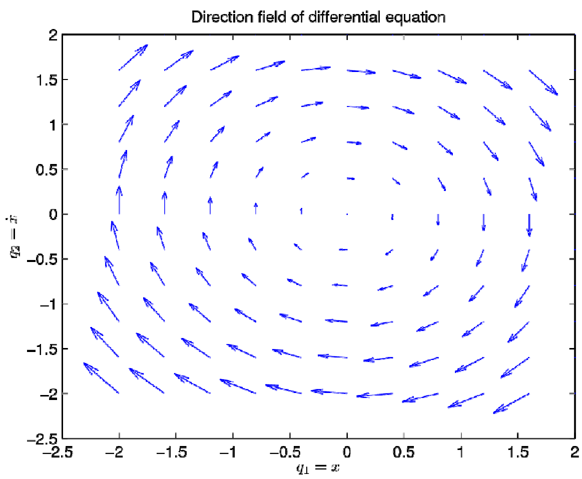
- Lecture Movie : http://www.youtube.com/watch?v=c0y5DcNQ8gs
- Lab : http://matrix.skku.ac.kr/knou-knowls/cla-week-11-sec-8-1.html

Many problems in science and engineering can be written as a mathematical problem of solving linear system of differential equations. In this section, we learn how to solve linear system of differential equations by using a matrix diagonalization.

Details can be found in the following websites:

http://www.math.psu.edu/tseng/class/Math251/Notes-LinearSystems.pdf

http://matrix.skku.ac.kr/CLAMC/chap8/Page83.htm
http://matrix.skku.ac.kr/CLAMC/chap8/Page84.htm
http://matrix.skku.ac.kr/CLAMC/chap8/Page85.htm
http://matrix.skku.ac.kr/CLAMC/chap8/Page86.htm
http://matrix.skku.ac.kr/CLAMC/chap8/Page87.htm



"It is through science that we prove, but through intuition that we discover. "

Jules Henri Poincaré (1854 - 1912)
http://en.wikipedia.org/wiki/Henri_Poincar%C3%A9

He is often described as a polymath, and in mathematics as The Last Universalist by Eric Temple Bell,[3] since he excelled in all fields of the discipline as it existed during his lifetime.

Chapter 8 Exercises

- http://matrix.skku.ac.kr/LA-Lab/index.htm
- http://matrix.skku.ac.kr/knou-knowls/cla-sage-reference.htm

Problem 1 Suppose a linear transformation T is defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ x_1 + x_2 \end{bmatrix}$ and $\alpha = \left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}\right\}$ is an ordered basis for $\mathbb{R}^2$. Find the matrix representation $[T]_\alpha$ of T relative to the ordered basis α.

Problem 2 Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be defined by $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x - 2y \\ 5x - y \\ 2x + 3y \end{bmatrix}$ and let $\alpha = \{\mathrm{v}_1, \mathrm{v}_2\}$, $\beta = \{\mathrm{v}_1', \mathrm{v}_2', \mathrm{v}_3'\}$ be ordered bases for R^2, R^3, respectively, where $\mathrm{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathrm{v}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathrm{v}_1' = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, $\mathrm{v}_2' = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\mathrm{v}_3' = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. Find the matrix representation $[T]_\alpha^\beta$ of T with respect to the ordered bases α and β.

Problem 3 Suppose a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -5x + 6y \\ -3x + 4y \end{bmatrix}$ and $\alpha = \left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$, $\beta = \left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$ are ordered bases for R^2.

(1) Find the matrix representation $[T]_\alpha$ of T relative to the ordered basis α.

(2) Find the transition matrix $P = [I]_\beta^\alpha$ from β to α.

(3) Compute $P^{-1}[T]_\alpha P$.

Solution (1) $[T]_\alpha = [T]_\alpha^\alpha = [[T(e_1)]_\alpha, [T(e_2)]_\alpha]$

$\Rightarrow T(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} -5 \\ -3 \end{bmatrix}$, $T(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$

$\Rightarrow [T]_\alpha = \begin{bmatrix} -5 & 6 \\ -3 & 4 \end{bmatrix}$

(2) $P = [I]_\alpha^\beta = [[\mathbf{y}_1]_\alpha, [\mathbf{y}_2]_\alpha]$

$\Rightarrow$ $a_1\begin{bmatrix}1\\0\end{bmatrix} + a_2\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix}, a_1 = 1, a_2 = 1,$

$b_1\begin{bmatrix}1\\0\end{bmatrix} + b_2\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}2\\1\end{bmatrix}, b_1 = 2, b_2 = 1,$

$\Rightarrow$ $P = \begin{bmatrix}a_1 & b_1\\ a_2 & b_2\end{bmatrix} = \begin{bmatrix}1 & 2\\ 1 & 1\end{bmatrix}$

(3) $P^{-1}[T]_\alpha P = \begin{bmatrix}-1 & 4\\ 1 & -1\end{bmatrix}\begin{bmatrix}-5 & 6\\ -3 & 4\end{bmatrix}\begin{bmatrix}1 & 2\\ 1 & 1\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & -2\end{bmatrix}$ ■

Problem 4 Determine if the given matrix A is diagonalizable. If A is diagonalizable, find matrix P diagonalizing A and the associated diagonal matrix D such that $D = P^{-1}AP$.

(1) $A = \begin{bmatrix}3 & -4\\ -4 & 3\end{bmatrix}$.

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

Solution Sage:

```
A=matrix([[2,1,1],[1,2,1],[1,1,2]])
print A.eigenvectors_right()
```

```
[(4, [(1, 1, 1)], 1), (1, [(1, 0, -1),(0, 1, -1)], 2)]
```

```
x1=vector([1, 1, 1])
x2=vector([1, 0, -1])
x3=vector([0, 1, -1])
P=column_matrix([x1,x2,x3])
print P
print P.det()
```

```
[ 1  1  0]
[ 1  0  1]
[ 1 -1 -1]
3
```
■

Problem 5 Find the algebraic and geometric multiplicity of each eigenvalue of A:

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

Problem 6 Find matrix P orthogonally diagonalizing matrix A and the diagonal matrix D such that $P^TAP=D$, using Sage.

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

Solution Sage:

```
A=matrix(QQ,3,3,[1,2,2,2,1,-2,2,-2,1])
A.eigenvalues()
```

[-3, 3, 3]

```
A=matrix(QQ,3,3,[1,2,2,2,1,-2,2,-2,1])
A.eigenvevtors_right()
```

[(-3, [(1, -1, -1)], 1), (3, [(1, 0, 1),(0, 1, -1)], 2)]

```
C=matrix(3,3,[0,2/sqrt(6),1/sqrt(3),1/sqrt(2),1/sqrt(6),-1/sqrt(3),-1/sqrt(2),1/s
qrt(6),-1/sqrt(3)])
C.transpose()*C
```

[1 0 0]

[0 1 0]

[0 0 1] ■

Problem 7 In each of the following matrix A and set S of linear independent eigenvectors of A are given. Find an orthogonal matrix P and a diagonal matrix D such that $P^TAP=D$.

(1) $A=\begin{bmatrix}-3&6\\6&-3\end{bmatrix}$, $S=\{(-2,2),(5,5)\}$.

(2) $A=\begin{bmatrix}7&2\\2&4\end{bmatrix}$, $S=\{(2,1),(1,-2)\}$.

Problem 8 Compute $q(\mathrm{x}) = \mathrm{x}^T A \mathrm{x}$ when

$$A = \begin{bmatrix} 3 & 2 & 1 \\ -1 & 0 & -4 \\ 5 & -2 & 1 \end{bmatrix},\ \mathrm{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Solution $q(x,y,z) = \mathrm{x}^T A\, \mathrm{x} = [x\ y\ z] \begin{bmatrix} 3 & 2 & 1 \\ -1 & 0 & -4 \\ 5 & -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 3x^2 + z^2 + xy + 6xz - 6yz.$ ■

Problem 9 Write the following expression as a quadratic form $\mathrm{x}^T A \mathrm{x}$:

$$2x^2 + 3y^2 + z^2 + xy - 2xz + 3yz.$$

Solution $q(x,y,z) = \mathrm{x}^T A\, \mathrm{x} = [x\ y\ z] \begin{bmatrix} 2 & 1/2 & -1 \\ 1/2 & 3 & 3/2 \\ -1 & 3/2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

Sage :
```
------------------------------------------
var('x','y','z')
a=2;    b=3;    c=1;    d=1;    e=-2;    f=3
A=matrix(3,3,[a,d/2,e/2,d/2,b,f/2,e/2,f/2,c])
X=matrix(3,1,[x,y,z])
print expand(X.transpose() * A * X)
------------------------------------------
[2*x^2 + x*y - 2*x*z + 3*y^2 + 3*y*z + z^2].
```
■

Problem 10 Eliminate the cross-product term from the following:

$$q(x,\ y) = x^2 - 2xy + y^2.$$

Problem 11 Sketch the graph of the following equation:

$$x^2 + 4xy + 4y^2 + 6x + 2y - 25 = 0$$

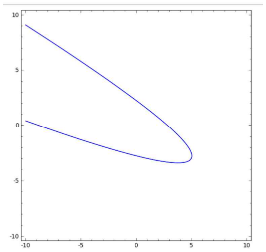
Solution $\mathrm{x}^T A \mathrm{x} + B\mathrm{x} - 25 = 0$ where $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, B = [6\ 2], \mathrm{x} = \begin{bmatrix} x \\ y \end{bmatrix}$.

=> $\lambda_1 = 5, \lambda_2 = 0$ => $\mathrm{v}_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathrm{v}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

=> $P = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$. $\mathbf{x} = P\mathbf{x}'$ and $5x'^2 + 2\sqrt{5}x' + 2\sqrt{5}y' - 25 = 0$.

=> $y' = \frac{1}{2\sqrt{5}}(26 - (\sqrt{5}x' + 1)^2)$

Sage :
```
------------------------------------------
var('x y')
f=x^2+4*x*y+4*y^2+6*x+2*y-25
implicit_plot(f==0, (x,-10,10), (y,-10,10))
------------------------------------------
```



■

Problem 12 Eliminate the cross-product terms from the quadratic surface $5x^2 + 6y^2 + 7z^2 + 4xy + 4yz = 1$ by properly rotating the axes.

Problem 13 Compute the singular values of matrix A:

$$A = [1 \ 2 \ 0].$$

Problem 14 Find the SVD of A:

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

Solution $AA^T = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $A^TA = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

Sage :
```
------------------------------------------
aat = matrix(QQ,2,2,[2,1,1,2])
ata = matrix(QQ,3,3,[1,1,0,1,2,1,0,1,1])
print aat.right_eigenvectors()
print ata.right_eigenvectors()
------------------------------------------
[(3, [(1, 1)], 1), (1, [(1, -1)], 1)]
[(3, [(1, 2, 1)], 1), (1, [(1, 0, -1)], 1), (0, [(1, -1, 1)], 1)]
```

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix}, \quad V = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix},$$

=> $$A = U\Sigma V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$ ■

Problem 15 In the below the SVD of A is given. Find $A^{\dagger}$.

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

Problem 16 The following matrix A has full column rank. Find its pseudo-inverse.

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 3 & 7 \end{bmatrix}.$$

Solution Sage :

```
------------------------------------------
A=matrix(QQ,[[1,1],[0,2],[3,7]])
print A.rank()
B=A.transpose() * A
Pseudo=B^-1*A.transpose()
print Pseudo
------------------------------------------
2
[    4/7  -11/14     1/7]
[  -3/14    5/14    1/14]
```

$$\therefore A^{\dagger} = (A^TA)^{-1}A^T = \begin{bmatrix} \frac{4}{7} & \frac{-11}{14} & \frac{1}{7} \\ \frac{-3}{14} & \frac{5}{14} & \frac{1}{14} \end{bmatrix}.$$ ■

Problem 17 For given vectors $\mathbf{u} = (2i, 0, 3i)$, $\mathbf{v} = (2-i, 0, 1+3i)$, compute Euclidean inner products $\mathbf{u} \cdot \mathbf{v}$ and $\mathbf{v} \cdot \mathbf{u}$.

Problem 18 Let $\mathrm{u}=\begin{bmatrix} 62-85i \\ 9 \\ 66i \\ 20i+6 \\ 63+11i \end{bmatrix}$, $\mathrm{v}=\begin{bmatrix} 31-55i \\ 13-61i \\ 21-63i \\ 11i \\ -71+9i \end{bmatrix}$ be vectors in C^5 with Euclidean inner product defined. Compute the norms $||\mathrm{u}||$ and $||\mathrm{v}||$, and $d(\mathrm{u}, \mathrm{v})$.

Problem 19 Find the eigenvalues of $A=\begin{bmatrix} 8 & 7 \\ 1 & -2 \end{bmatrix}$ and a basis for the eigenspace associated with each eigenvalue.

Problem 20 Find any invertible matrix P diagonalizing a given matrix A which has complex eigenvalues?

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

Solution Sage : --

```
A=matrix(QQ,[[6,-4],[8,-2]])
print A.eigenvalues()
print A.eigenvectors_right()
```

[2 - 4*I, 2 + 4*I][(2 - 4*I, [(1, 1 + 1*I)], 1), (2 + 4*I, [(1, 1 - 1*I)], 1)].

$$\therefore\ P=\begin{bmatrix} 1 & 1 \\ 1+i & 1-i \end{bmatrix}$$ ■

Problem 21 Find the conjugate transpose A^* of the following matrix A:

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

Problem 22 Determine which matrices in the below are Hermitian.

(a) $\begin{bmatrix} 3+i & 2 \\ 7 & 4-i \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2+i \\ 2-i & -1 \end{bmatrix}$

(c) $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ (d) $\begin{bmatrix} \frac{-1}{\sqrt{2}}i & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}i \end{bmatrix}$

(e) $\begin{bmatrix} 0 & i & 1 \\ i & 0 & 2+i \\ -1 & -2+i & 0 \end{bmatrix}$ (f) $\begin{bmatrix} 2 & i & 0 \\ 0 & 1 & -5i \\ 1 & 1-i & 4 \end{bmatrix}$

Problem 23 Determine if each matrix in the below is unitary.

(a) $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$ (b) $A = \begin{bmatrix} 0 & i & 1 \\ i & 0 & 2+i \\ -1 & -2+i & 0 \end{bmatrix}$

Problem 24 Replace each $\times$ by a complex number to make matrix A Hermitian.

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

Problem 25 Show that the following matrix A is unitary, and find its inverse A^{-1}.

$$A = \begin{bmatrix} \frac{1}{2\sqrt{2}}(\sqrt{3}+i) & \frac{1}{2\sqrt{2}}(1-i\sqrt{3}) \\ \frac{1}{2\sqrt{2}}(1+i\sqrt{3}) & \frac{1}{2\sqrt{2}}(i-\sqrt{3}) \end{bmatrix}$$

Problem P1 Let $v_1 = (2, 0)$ and $v_2 = (0, 4)$ and suppose $T : R^2 \to R^2$ is a linear transformation. If $T(v_1) = v_2$ and $T(v_2) = v_1$, what is the standard matrix of T? In addition, if $\beta = \{v_1, v_2\}$, what is the matrix representation $[T]_\beta$ of T relative to the ordered basis β?

Problem P2 Suppose the following polynomial $p(\lambda)$ is the characteristic polynomial of a square matrix A.

$p(\lambda) = (\lambda - 1)(\lambda - 3)^2(\lambda - 4)^3$.

(1) What is the order of A?

(2) If the number of linear independent eigenvectors of A cannot exceed 3, is the matrix A diagonalizable?

(3) What is the dimension of each eigenspace of A?

(4) Suppose A is diagonalizable. Discuss about a relationship between the algebraic multiplicity of each eigenvalue λ and the dimension of the solution space to the homogeneous linear system $(\lambda I - A)\mathbf{x} = \mathbf{0}$.

Problem P3 (1) Suppose the following are the eigenvalues of a 3×3 symmetric matrix A and their corresponding eigenvectors:

$\lambda_1 = -1,\ \lambda_2 = 3,\ \lambda_3 = 7 \qquad \mathbf{v}_1 = (0,\ 1,\ -1),\ \mathbf{v}_2 = (1,\ 0,\ 0),\ \mathbf{v}_3 = (0,\ 1,\ 1)$.

Find the matrix A.

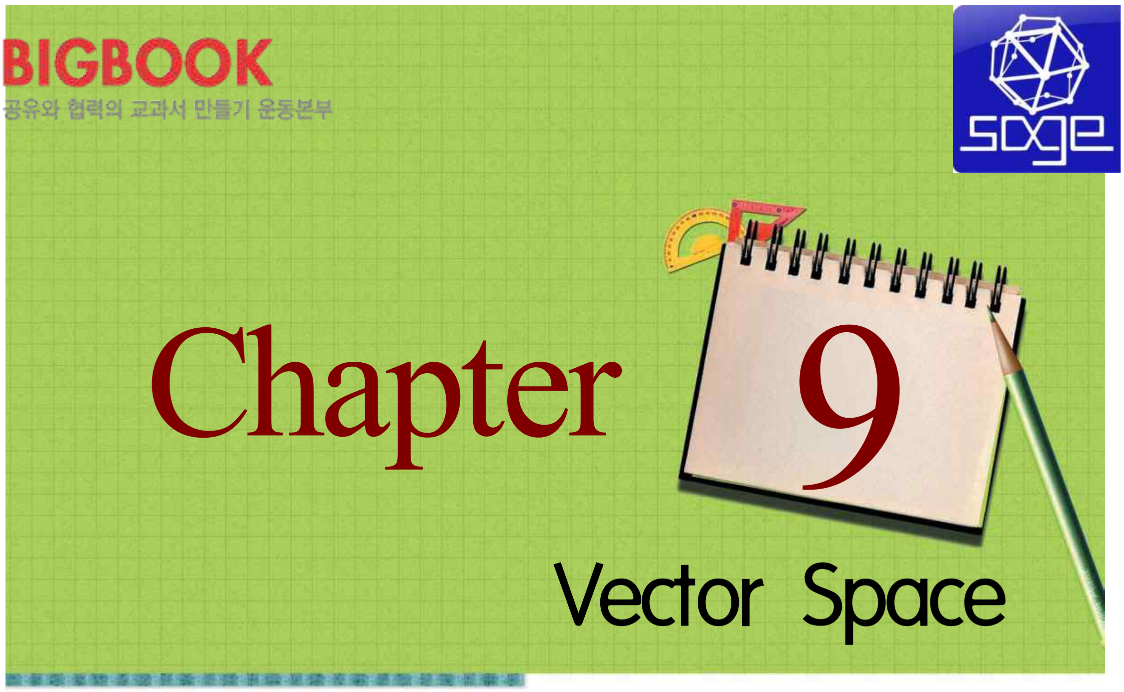
(2) Determine if there exists a 3×3 matrix whose eigenvalues and their corresponding eigenvectors are given in the below:

$$\lambda_1 = -1,\ \lambda_2 = 3,\ \lambda_3 = 7 \qquad \mathbf{v}_1 = (0,\ 1,\ -1),\ \mathbf{v}_2 = (1,\ 0,\ 0),\ \mathbf{v}_3 = (1,\ 1,\ 1)$$

Problem P4 Show $A = \begin{bmatrix} 12 & 6i \\ 6i & 10 \end{bmatrix}$ has non-real eigenvalues.

Problem P5 Show that if $A \in M_n(C)$ is skew-Hermitian, then every eigenvalue of A is a pure imaginary number.

Problem P6 Use properties (1) and (3) of inner product in Section 8.7 to show that $< c\mathbf{u}, \mathbf{v} > = \bar{c} < \mathbf{u}, \mathbf{v} >$.



The operations used in vector addition and scalar multiple are not limited to the theory but can be applied to all areas in society. For example, consider objects around you as vectors and make a set of vectors, then create two proper operations (vector addition and scalar multiple) from the relations between the objects. If these two operations satisfy the two basic laws and 8 operation properties, the set becomes a mathematical vector space (or linear space). Thus we can use all properties of a vector space and can analyze the set theoretically and apply them to real problems.

In this chapter, we give a definition of a vector space and a general theory of a vector space.

9.1 Axioms of a Vector Space

- Ref site : http://youtu.be/m9ru-F7EvNg, http://youtu.be/beXWYXYtAal
- Lab site: http://matrix.skku.ac.kr/knou-knowls/cla-week-14-sec-9-1.html

The concept of vectors has been extended to n-tuples in $\mathbb{R}^n$ from the arrows in the 2-dimensional or 3-dimensional space. In Chapter 1, we defined the addition and the scalar multiple in the n-dimensional space $\mathbb{R}^n$. In this section, we extend the concept of the n-dimensional space $\mathbb{R}^n$ to an n-dimensional vector space.

Vector Spaces

Definition [Vector space]

If a set $V(\neq \phi)$ has two well-defined binary operations, vector addition (A) '$+$' and scalar multiplication (SM) '$\cdot$', and for any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and $h, k \in \mathbb{R}$, two basic laws

A. $\mathbf{x}, \mathbf{y} \in V \Rightarrow \mathbf{x}+\mathbf{y} \in V$.
SM. $\mathbf{x} \in V, k \in \mathbb{R} \Rightarrow k\mathbf{x} \in V$.

and the following eight laws hold, then we say that the set V forms a vector space over $\mathbb{R}$ with the given two operations, and we denote it by $(V, +, \cdot)$ (simply V if there is no confusion). Elements of V are called vectors.

A1. $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$.
A2. $(\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z})$.
A3. For any $\mathbf{x} \in V$, there exists a unique element $\mathbf{0}$ in V such that $\mathbf{x}+\mathbf{0} = \mathbf{x}$.
A4. For each element $\mathbf{x}$ of V, there exists a unique $-\mathbf{x}$ such that $\mathbf{x}+(-\mathbf{x}) = \mathbf{0}$.

SM1. $k(\mathbf{x}+\mathbf{y}) = k\mathbf{x}+k\mathbf{y}$.
SM2. $(h+k)\mathbf{x} = h\mathbf{x}+k\mathbf{x}$.
SM3. $(hk)\mathbf{x} = h(k\mathbf{x}) = k(h\mathbf{x})$.
SM4. $1\mathbf{x} = \mathbf{x}$.

The vector $\mathbf{0}$ satisfying A3 is called a zero vector, and the vector $-\mathbf{x}$ satisfying A4 is called a negative vector of $\mathbf{x}$.

- In general, the two operations defining a vector space are important. Therefore, it is better to write $(V, +, \cdot)$ instead of just V.

Example 1

For vectors $\mathbf{x} = (x_1, x_2, x_3)$, $\mathbf{y} = (y_1, y_2, y_3)$ in $\mathbb{R}^3$ and a scalar $k \in \mathbb{R}$, the vector sum $\mathbf{x}+\mathbf{y}$ and a scalar multiple $k\mathbf{x}$ by $k \in \mathbb{R}$ are defined as

(1) $\mathbf{x}+\mathbf{y} = (x_1+y_1, x_2+y_2, x_3+y_3)$.

(2) $k\mathbf{x} = (kx_1, kx_2, kx_3)$.

The set $(\mathbb{R}^3, +, \cdot)$ together with the above operations forms a vector space over the set $\mathbb{R}$ of real numbers.

■

Example 2

For vectors in $\mathbb{R}^n$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

and a scalar $k \in \mathbb{R}$, the sum of two vectors $\mathbf{x}+\mathbf{y}$ and the scalar multiple of $\mathbf{x}$ by k is defined by

$$(1)\ \mathbf{x}+\mathbf{y} = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix} \quad \text{and} \quad (2)\ k\mathbf{x} = \begin{bmatrix} kx_1 \\ kx_2 \\ \vdots \\ kx_n \end{bmatrix}.$$

The set $\mathbb{R}^n$ form a vector space together with the above two operations.

■

Theorem 9.1.1

Let V be a vector space. Let $\mathbf{x} \in V$ and $k \in \mathbb{R}$. Then the following hold.

(1) $0\mathbf{x} = \mathbf{0}$.

(2) $k\mathbf{0} = \mathbf{0}$.

(3) $(-1)\mathbf{x} = -\mathbf{x}$.

(4) $k\mathbf{x} = \mathbf{0} \quad \Leftrightarrow \quad k = 0$ or $\mathbf{x} = \mathbf{0}$.

Zeo Vector Space

Definition

Let $V = \{\mathbf{0}\}$. For a scalar $k \in \mathbb{R}$, if the addition and scalar multiple are defined as

$$\mathbf{0} + \mathbf{0} = \mathbf{0}, \quad k\mathbf{0} = \mathbf{0}, \text{ then}$$

V forms a vector space. This vector space is called a zero vector space.

Example 3 Let $M_{m \times n}$ be the set of all $m \times n$ matrices with real entries. That is,

$$M_{m \times n} = \{A = [a_{ij}]_{m \times n} | a_{ij} \in \mathbb{R},\ 1 \le i \le m,\ 1 \le j \le n\}.$$

When $m = n$, we denote $M_{m \times n}$ by M_n.

If $M_{m \times n}$ is equipped with the matrix addition and the scalar multiplication, then $M_{m \times n}$ form a vector space $(M_{m \times n}, +, \cdot\ )$ over $\mathbb{R}$. Then the zero vector is the zero matrix O and for each $A = [a_{ij}] \in M_{m \times n}$, the negative vector is $-A = [-a_{ij}]$. Note that each vector means an $m \times n$ matrix with real entries. ■

Example **4** Let $\mathcal{E}(\mathbb{R})$ be the set of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$. That is,

$$\mathcal{E}(\mathbb{R}) = \{f \mid f : \mathbb{R} \to \mathbb{R} \text{ is continuous}\}$$

Let $f,\ g \in \mathcal{E}(\mathbb{R})$ and a scalar $k \in \mathbb{R}$, define the addition and the scalar multiple as

$$(f+g)(x) = f(x)+g(x),\ (kf)(x) = kf(x).$$

Then $\mathcal{E}(\mathbb{R})$ forms a vector space $(\mathcal{E}(\mathbb{R}),\ +, \cdot\ )$ over $\mathbb{R}$.

Now the zero vector is $\mathbf{0}(x) = 0$ and for each $f \in \mathcal{E}(\mathbb{R})$, $-f$ is defined as $(-f)(x) = -f(x)$.

Vectors in $\mathcal{E}(\mathbb{R})$ mean continuous functions from $\mathbb{R}$ to $\mathbb{R}$. ■

Example **5** Let P_n be the set of all polynomials of degree at most n with real coefficients. In other words,

$$P_n = \left\{a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n \mid a_0,\ a_1,\ \ldots,\ a_n \in \mathbb{R}\right\}$$

Let $p(t) = a_0 + a_1 t + \cdots + a_n t^n,\ q(t) = b_0 + b_1 t + \cdots + b_n t^n \in P_n$ and a scalar $k \in \mathbb{R}$. The addition and the scalar multiplication are defined as

$$p(t)+q(t) = (a_0+b_0)+(a_1+b_1)t+\ \cdots\ +(a_n+b_n)t^n,$$

$$kp(t) = (ka_0)+(ka_1)t+\ \cdots\ +(ka_n)t^n.$$

Then P_n forms a vector space $(P_n,\ +, \cdot\ )$ over $\mathbb{R}$. Now the zero vector is $\mathbf{0}(t) = 0 + 0t +\ \cdots\ + 0t^n$ and each $p(t) \in P_n$ has the negative vector $-p(t)$ defined as

$$-p(t) = -a_0 - a_1 t -\ \cdots\ - a_n t^n.$$

Vectors in P_n means polynomials of degree at most n with real coefficients. ■

Subspace

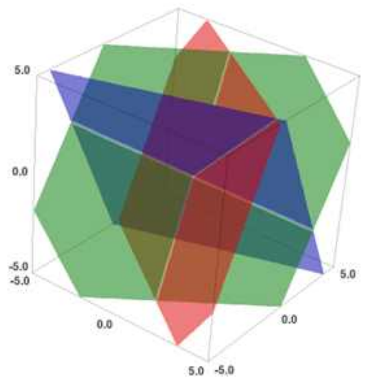
Definition

Let V be a vector space and $W(\neq \phi)$ be a subset of V. If W forms a vector space with the operations defined in V, then W is called a subspace of V.

Example 6 If $(V, +, \cdot\ )$ is a vector space, $\{\mathbf{0}\}$ and V itself are subspaces of V. ■

- In fact, the only subspaces of R^2 are $\{0\}$, R^2, and lines passing through the origin. (see section 3.4 Example 3).

- In $\mathbb{R}^3$, only subspaces are (i) Null Spaces, (ii) $\mathbb{R}^3$, (iii) lines passing through origin and (iv) planes passing through origins.



How to determine a subspace? (the 2-step subspace test)

Theorem 9.1.2 [the 2−step subspace test]

Let a set $(V, +, \cdot\ )$ be a vector space and $W(\neq \varnothing)$ be a subset of V. A necessary and sufficient condition for W to be a subspace of V is

(1) $\mathbf{x},\ \mathbf{y} \in W \Rightarrow \mathbf{x}+\mathbf{y} \in W$ (closed under vector addition $+$)

(2) $\mathbf{x} \in W,\ k \in \mathbb{R} \Rightarrow k\mathbf{x} \in W$ (closed under scalar multiple $\cdot$)

Example **7**

Show that $W = \left\{ \begin{bmatrix} 0 & a & b \\ c & d & 0 \end{bmatrix} \,\middle|\, a, b, c, d \in \mathbb{R} \right\}$ is a subspace of the vector space $(M_{2\times 3}, +, \cdot\ )$.

Solution

Note that $M_{2\times 3}$ is a vector space under the matrix addition and the scalar multiplication. Let

$$\mathbf{x} = \begin{bmatrix} 0 & a_1 & b_1 \\ c_1 & d_1 & 0 \end{bmatrix},\ \mathbf{y} = \begin{bmatrix} 0 & a_2 & b_2 \\ c_2 & d_2 & 0 \end{bmatrix} \in W,\ k \in \mathbb{R}.$$

The following two conditions are satisfied.

(1) $\mathbf{x} + \mathbf{y} = \begin{bmatrix} 0 & a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 & 0 \end{bmatrix} \in W,$

(2) $k\mathbf{x} = \begin{bmatrix} 0 & ka_1 & kb_1 \\ kc_1 & kd_1 & 0 \end{bmatrix} \in W.$

Hence by Theorem **9.1.2**, $(W +, \cdot\ )$ is a subspace of $(M_{2\times 3}, +, \cdot\ )$. ■

Example **8**

The set of invertible matrices of order n is not a subspace of the vector space M_n.

Solution

One can make a non-invertible matrix by adding two invertible matrices. For example,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$ ■

Example **9**

Let V be a vector space and $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_t\} \subseteq V$. Show that the set

$$W = \{c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_t\mathbf{x}_t \mid c_1, c_2, \ldots, c_t \in \mathbb{R}\}$$

is a subspace of V. Note that $W = \langle S \rangle$, linear span of the set S.

Solution

Suppose that $\mathbf{x}, \mathbf{y} \in W$, $k \in \mathbb{R}$. Then for $c_i, d_i \in \mathbb{R}$ $(i = 1, 2, \ldots, t)$,

$$\mathrm{x} = c_1\mathrm{x}_1 + c_2\mathrm{x}_2 + \cdots + c_t\mathrm{x}_t,\ \ \mathrm{y} = d_1\mathrm{x}_1 + d_2\mathrm{x}_2 + \cdots + d_t\mathrm{x}_t\ .$$

Thus

$$\mathrm{x} + \mathrm{y} = (c_1 + d_1)\mathrm{x}_1 + (c_2 + d_2)\mathrm{x}_2 + \cdots + (c_t + d_t)\mathrm{x}_t\ ,$$
$$k\mathrm{x} = (kc_1)\mathrm{x}_1 + (kc_2)\mathrm{x}_2 + \cdots + (kc_t)\mathrm{x}_t\ .$$

$$\therefore\ \ \mathrm{x} + \mathrm{y} \in W,\quad k\mathrm{x} \in W$$

Therefore W is a subspace of V. ■

Liinear independence and linear dependence

Definition [Linear independence and linear dependence]

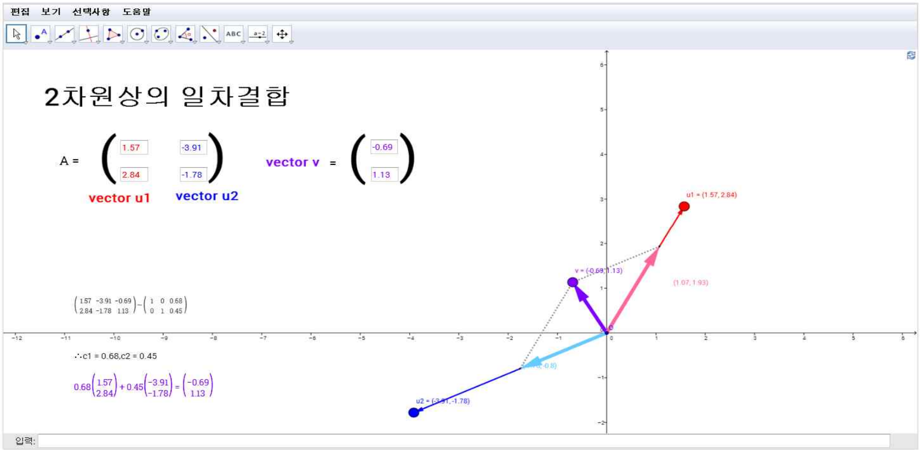
If a subset $\{\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_n\}$ of a vector space V satisfies the following condition, it is called linearly independent.

$$c_1\mathrm{v}_1 + c_2\mathrm{v}_2 + \cdots + c_n\mathrm{v}_n = \mathbf{0} \quad \Rightarrow \quad c_1 = c_2 = \cdots = c_n = 0$$

and if the set is not linearly independent, it is called linearly dependent. Hence being linearly independent means that there exist some scalars $c_1, c_2, \ldots, c_n$ not all zero such that $c_1\mathrm{v}_1 + c_2\mathrm{v}_2 + \cdots + c_n\mathrm{v}_n = \mathbf{0}$.

Remark Linear combination in 2-dimensional space - linear dependence (computer simulation)

- http://www.geogebratube.org/student/m57551



Example 10 Let $E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Since

$$c_1E_{11} + c_2E_{12} + c_3E_{21} + c_4E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \;\Rightarrow\; c_1 = c_2 = c_3 = c_4 = 0$$

$\{E_{11}, E_{12}, E_{21}, E_{22}\}$ is a linearly independent set of M_2. ■

Example 11 Let $A = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$, $C = \begin{bmatrix} 0 & -1 \\ 2 & 2 \end{bmatrix}$. Since $A = B + C$, $\{A, B, C\}$ is a linearly dependent set of M_2. ■

Example 12 The subset $\{1, x, x^2, \ldots, x^n\}$ of P_n is linearly independent. ■

Example 13 Let $\{2 - x + x^2, 2x + x^3, 4 - 4x + x^2\}$ be a subset of P_3. Then since

$$4 - 4x + x^2 = 2(2 - x + x^2) - (2x + x^2),$$

the set is linearly dependent. ■

Basis

Definition **[basis and dimension]**

If a subset $\alpha(\neq \phi)$ of a vector space V satisfies the following conditions, α is a **basis** of V.

(1) $\text{span}(\alpha) = V$.
(2) α is linearly independent.

In this case, the number of elements of the basis α, $|\alpha|$, is called the **dimension** of V, denoted by $\dim(V)$.

Example 14

The set in Example 10 consisting of $E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ is a basis of M_2. Thus $\dim(M_{2\times 2}) = 4$. On the other hand, the set in Example 12 $\{1, x, x^2, \ldots, x^n\}$ is a basis of P_n. Thus $\dim(P_n) = n+1$. These bases play a role similar to the standard basis of R^n, hence M_2 and P_n are called **standard bases.** ■

Example 15

Show that $\alpha = \{1+x, -1+x, x^2\}$ is a basis of P_2.

Solution

$$a(1+x)+b(-1+x)+cx^2 = 0 \Leftrightarrow (a-b)+(a+b)x+cx^2 = 0$$
$$\Leftrightarrow a-b=0,\ a+b=0,\ c=0$$

Since $a=b=c=0$, α is linearly independent.

Next, given $A+Bx+Cx^2 \in P_2$, the existence of a, b, c such that

$$A+Bx+Cx^2 = a(1+x)+b(-1+x)+cx^2$$

is guaranteed since the coefficient matrix of the linear system

$$\begin{cases} a-b &= A \\ a+b &= B \\ c &= C \end{cases} \quad \text{that is,} \quad \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

is invertible. Thus α spans P_2. Hence α is a basis of P_2. ■

Linear independence of continuous function: Wronskian

Theorem 9.1.3 [Wronski's Test]

If $f_1(x), f_2(x), \ldots, f_n(x)$ are $n-1$ times differentiable on the interval $(-\infty, \infty)$ and there exists $x_0 \in (-\infty, \infty)$ such that Wronskian $W(x_0)$ defined below is not zero, then these functions are linearly independent.

$$W(x_0) = \begin{vmatrix} f_1(x_0) & \cdots & f_n(x_0) \\ f_1{}'(x_0) & \cdots & f_n{}'(x_0) \\ \vdots & \vdots & \vdots \\ f_1^{(n-1)}(x_0) & \cdots & f_n^{(n-1)}(x_0) \end{vmatrix} \neq 0$$

Conversely if $W(x)=0$ for every x in $(-\infty, \infty)$, then $f_1, \ldots, f_n$ are linearly dependent.

Example 16 Show **by Theorem 9.1.3** that $f_1(x)=1$, $f_2(x)=e^x$, $f_3(x)=e^{2x}$ are linearly independent.

Solution

For some (in fact, any) x, $W(x) = \begin{vmatrix} 1 & e^x & e^{2x} \\ 0 & e^x & 2e^{2x} \\ 0 & e^x & 4e^{2x} \end{vmatrix} = 2e^{3x} \neq 0$. Thus these functions are linearly independent. □

Sage http://sage.skku.edu or http://mathlab.knou.ac.kr:8080/

```
var('x')
W=wronskian(1, e^x, e^(2*x))    # wronskian(f1(x), f2(x), f3(x))
print W
```

2*e^(3*x) ■

Example 17 Let $f_1(x)=x$, $f_2(x)=\sin x$. Show that these functions are linearly independent.

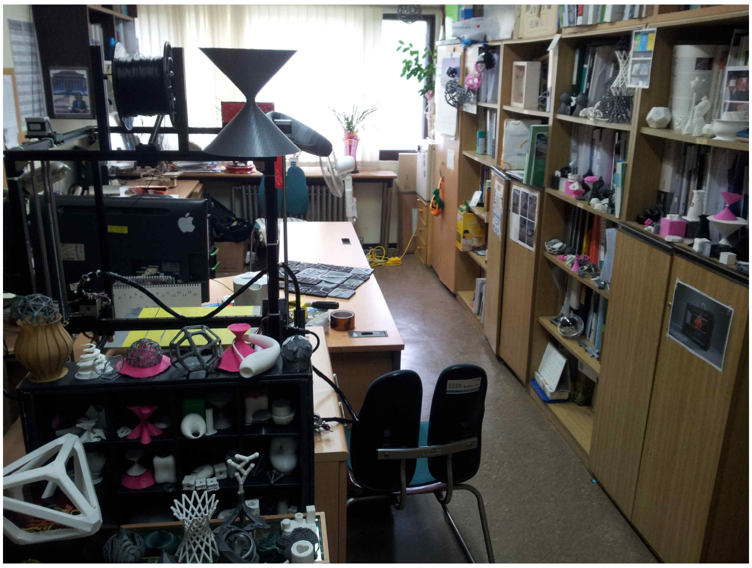
Solution

Since $W(x_0)=\begin{vmatrix} x_0 & \sin x_0 \\ 1 & \cos x_0 \end{vmatrix}=x_0\cos x_0-\sin x_0 \neq 0$ for some x_0, these functions are linearly independent. ■

Example 18 Show that $f_1(x)=x$, $f_2(x)=3x$ are linearly dependent.

Solution

Since for any x, $W(x)=\begin{vmatrix} x & 3x \\ 1 & 3 \end{vmatrix}=0$, these functions are linearly dependent. ■



http://matrix.skku.ac.kr/kiosk/

9.2 Inner product; *Fourier series

- ref movie: http://youtu.be/m9ru-F7EvNg, http://youtu.be/nIkYF-uvFdA
- demo site: http://matrix.skku.ac.kr/knou-knowls/cla-week-14-sec-9-2.html

In this section, we generalize the Euclidean inner product on $\mathbb{R}^n$ (dot product) to introduce the concepts of length, distance, and orthogonality in a general vector space.

Inner product and inner product space

Definition [Inner product and inner product space]

The inner product on a real vector space V is a function assigning a pair of vectors $\mathbf{u}$, $\mathbf{v}$ to a scalar $<\mathbf{u}, \mathbf{v}>$ satisfying the following conditions. (that is, the function $<\ ,\ >: V \times V \to \mathbb{R}$ satisfies the following conditions.)

(1) $<\mathbf{u}, \mathbf{v}> = <\mathbf{v}, \mathbf{u}>$ for every $\mathbf{u}, \mathbf{v}$ in V. f
(2) $<\mathbf{u}+\mathbf{v}, \mathbf{w}> = <\mathbf{u}, \mathbf{w}> + <\mathbf{v}, \mathbf{w}>$ for every $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in V.
(3) $<c\mathbf{u}, \mathbf{v}> = c<\mathbf{u}, \mathbf{v}>$ for every $\mathbf{u}, \mathbf{v}$ in V and c in $\mathbb{R}$.
(4) $<\mathbf{u}, \mathbf{u}> \geq 0$; $<\mathbf{u}, \mathbf{u}> = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}$ for every $\mathbf{u}$ in V.

The inner product space is a vector space V with an inner product $<\mathbf{u}, \mathbf{v}>$ defined on V.

Example 1 The Euclidean inner product, that is, the dot product is an example of an inner product on $\mathbb{R}^n$. Let us ask how other inner products on R^n are possible. For this, consider $A \in M_n(\mathbb{R})$. Let $\mathbf{u}$ and $\mathbf{v}$ be the column vectors of R^n. Define $<\mathbf{u}, \mathbf{v}>$ by $<\mathbf{u}, \mathbf{v}> = \mathbf{v}^T A\mathbf{u}$. Then let us find the condition on A so that this function becomes an inner product.

Solution

In order for $<\mathbf{u}, \mathbf{v}> = \mathbf{v}^T A\mathbf{u}$ to be an inner product, the four conditions (1)~(4) should be satisfied. First consider conditions (2) and (3).

$$
\begin{aligned}
<\mathbf{u}+\mathbf{v},\ \mathbf{w}> &= \mathbf{w}^T A(\mathbf{u}+\mathbf{v}) \\
&= \mathbf{w}^T A\mathbf{u} + \mathbf{w}^T A\mathbf{v} = <\mathbf{u},\ \mathbf{w}> + <\mathbf{v},\ \mathbf{w}>,
\end{aligned}
$$

$$<c\mathbf{u},\ \mathbf{v}> = \mathbf{v}^T A(c\mathbf{u}) = c\mathbf{v}^T A\mathbf{u} = c<\mathbf{u},\ \mathbf{v}>.$$

Let us check when condition (1) holds. Since $\mathbf{v}^T A\mathbf{u}$ is a 1×1 matrix (hence a real number), we have

$$(\mathbf{v}^T A\mathbf{u})^T = \mathbf{v}^T A\mathbf{u}$$

That is, to satisfy $<\mathbf{u},\ \mathbf{v}> = \mathbf{v}^T A\mathbf{u} = (\mathbf{v}^T A\mathbf{u})^T = \mathbf{u}^T A^T \mathbf{v} = \mathbf{u}^T A\mathbf{v} = <\mathbf{v},\ \mathbf{u}>$ we have $A = A^T$, in other words, A is a symmetric matrix.

Thus the function $<\mathbf{u},\ \mathbf{v}> = \mathbf{v}^T A\mathbf{u}$ satisfy condition (1) if A is a symmetric matrix.

Finally check condition (4). An $n\times n$ symmetric matrix A should satisfy $\mathbf{u}^T A\mathbf{u} > 0$ for any nonzero vector $\mathbf{u}$. This condition means that A is positive definite. In other words, if A is positive definite, $<\mathbf{u},\ \mathbf{v}> = \mathbf{v}^T A\mathbf{u}$ satisfies condition (4).

Therefore, to wrap up, if A is an $n\times n$ symmetric and positive definite matrix, then $<\mathbf{u},\ \mathbf{v}> = \mathbf{v}^T A\mathbf{u}$ defines an inner product on $\mathbb{R}^n$. The well known Euclidean inner product $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}^T\mathbf{u} = \mathbf{v}^T I\,\mathbf{u}$ can be obtained as a special case when $A = I_n$ (symmetric and positive definite). ■

- For any nonzero vector $\mathbf{u}$, if the eigenvalues of A are positive, then $\mathbf{u}^T A\mathbf{u} > 0$ (the converse also holds.)

Example 2

Let $A = \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix}$ be a 2×2 symmetric matrix and $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ in $\mathbb{R}^2$. Then

$$<\mathbf{u},\ \mathbf{v}> = \mathbf{v}^T A\mathbf{u} = 3v_1u_1 + 2v_2u_1 + 2v_1u_2 + 4v_2u_2$$

satisfies conditions (1), (2), (3) of an inner product on $\mathbb{R}^2$. Now let us show that A is a positive definite. Let $\mathbf{w} = \begin{bmatrix} x \\ y \end{bmatrix}$. Then

$< \mathrm{w}, \mathrm{w}> = \mathrm{w}^T A \mathrm{w} = 3x^2 + 4xy + 4y^2 = (x+2y)^2 + 2x^2$. Thus $\mathrm{w}^T A \mathrm{w} \geq 0$ and

$$\mathrm{w}^T A \mathrm{w} = 0 \Leftrightarrow x + 2y = 0 = x \Leftrightarrow x = y = 0.$$

Hence the symmetric matrix A is positive definite and defines an inner product on R^2 of the form $< \mathrm{u}, \mathrm{v}> = \mathrm{v}^T A \mathrm{u}$.

If $\mathrm{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\mathrm{v} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$, then $\mathrm{u} \cdot \mathrm{v} = 7$. On the other hand,

$$< \mathrm{u}, \mathrm{v}> = [1\ 4] \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 51$$

Hence the inner product $< \mathrm{u}, \mathrm{v}> = \mathrm{v}^T A \mathrm{u}$ on $\mathbb{R}^2$ is different from the Euclidean inner product. ■

Norm and angle

Definition [norm and angle]

Let V be a vector space with an inner product $< \mathrm{u}, \mathrm{v}>$. The norm (or length) of a vector u with respect to the inner product is defined by $\|\mathrm{u}\| = \sqrt{< \mathrm{u}, \mathrm{u}>}$.

The **angle** θ between two nonzero vectors u and v is defined by

$$\cos\theta = \frac{< \mathrm{u}, \mathrm{v}>}{\|\mathrm{u}\| \|\mathrm{v}\|} \quad (0 \leq \theta \leq \pi).$$

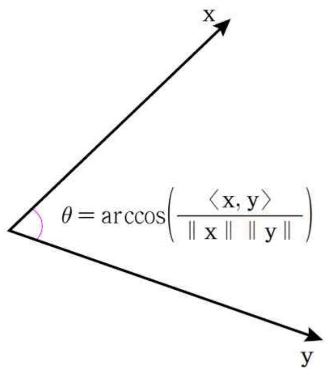
In particular, if two vectors u and v satisfy $< \mathrm{u}, \mathrm{v}> = 0$, then they are said to be orthogonal.

- For example, the norm of $\mathrm{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ with respect to the inner product given in Example 2 is

$$\|\mathrm{u}\|^2 = < \mathrm{u}, \mathrm{u}> = \mathrm{u}^T A \mathrm{u} = [3\ \ 1] \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 43.$$

Thus $\|\mathrm{u}\| = \sqrt{43}$. On the other hand, the norm $\mathrm{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ with respect to the Euclidean inner product is

$$\|\mathrm{u}\| = \sqrt{\mathrm{u} \cdot \mathrm{u}} = \mathrm{u}^T \mathrm{u} = \sqrt{10}$$



- For any inner prodcut space, the triangle inequality $\|\mathrm{u}+\mathrm{v}\| \leq \|\mathrm{u}\| + \|\mathrm{v}\|$ holds.

- Using the Gram-Schmidt orthogonality process, we can make a basis $\{\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_n\}$ of a inner product space V into an orthonormal basis $\{\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_n\}$.

Inner product on complex vector space

Definition

Let V be a complex vector space. Let $\mathrm{u}, \mathrm{v}, \mathrm{w}$ be any vectors in V and $c \in C$ be any scalar. The function $<,\ >$ from $V \times V$ to C is called an inner product (or Hermitian inner product) if the following hold.

(1) $< \mathrm{u}, \mathrm{v} > = < \overline{\mathrm{v}, \mathrm{u}} >$.

(2) $< \mathrm{u}+\mathrm{v}, \mathrm{w} > = < \mathrm{u}, \mathrm{w} > + < \mathrm{v}, \mathrm{w} >$.

(3) $< c\mathrm{u}, \mathrm{v} > = c < \mathrm{u}, \mathrm{v} >$.

(4) $< \mathrm{v}, \mathrm{v} > \geq 0$, $< \mathrm{v}, \mathrm{v} > = 0 \Leftrightarrow \mathrm{v} = 0$

- A complex vector space with an inner product is called a complex inner product space or a unitary space. If $< \mathrm{u}, \mathrm{v} > = 0$ for any two nonzero vectors u, v, then we say that u and v are orthogonal.

- Let V be a complex vector space. By the definition of an inner product on V, we obtain the following properties.

(1) $<0, \mathrm{v}> \; = 0 = \; <\mathrm{v}, 0>$.

(2) $<\mathrm{u}, \mathrm{v}+\mathrm{w}> \; = \; <\mathrm{u}, \mathrm{v}> + <\mathrm{u}, \mathrm{w}>$.

(3) $<\mathrm{u}, c\mathrm{v}> \; = \bar{c} <\mathrm{u}, \mathrm{v}>$, $c \in C$ $\quad (\because \; <\mathrm{u}, c\mathrm{v}> \; = (c\mathrm{v})^* A\mathrm{u} = \bar{c}\,\mathrm{v}^* A\mathrm{u} = \bar{c} <\mathrm{u}, \mathrm{v}>)$.

Example 3

Let $\mathrm{u} = (u_1, u_2, \ldots, u_n)$ and $\mathrm{v} = (v_1, v_2, \ldots, v_n)$ be vectors in C^n. The Euclidean inner product $\mathrm{u} \cdot \mathrm{v} = \overline{v_1} u_1 + \overline{v_2} u_2 + \cdots + \overline{v_n} u_n$ satisfies the conditions (1)~(4) for the inner product. ■

Example 4

Let $\mathcal{E}([a, b], C)$ be the set of continuous functions from the interval $[a, b]$ to the complex set C. Let $\boldsymbol{f}(x), \boldsymbol{g}(x) \in \mathcal{E}([a, b], C)$. If the addition and scalar multiple of these functions are defined below, then $\mathcal{E}([a, b], C)$ is a complex vector space with respect to these operations.

$$(\boldsymbol{f}+\boldsymbol{g})(x) = \boldsymbol{f}(x) + \boldsymbol{g}(x), \quad (c\boldsymbol{f})(x) = c\boldsymbol{f}(x), \; c \in C.$$

In this case, a vector in $\mathcal{E}([a, b], C)$ is of the form $\boldsymbol{f}(x) = f_1(x) + i f_2(x)$ and $f_1(x), f_2(x)$ are continuous functions from $[a, b]$ to $\mathbb{R}$. For $\boldsymbol{f}(x)$, $\boldsymbol{g}(x) \in \mathcal{E}([a, b], C)$, define the following inner product

$$<\boldsymbol{f}(x), \boldsymbol{g}(x)> = \int_a^b \overline{\boldsymbol{g}(x)}\boldsymbol{f}(x)\, dx.$$

Then $\mathcal{E}([a, b], C)$ is a complex inner product space.

We leave readers to check conditions (1)~(3) for an inner product, and show condition (4) here. Note

$$<\boldsymbol{f}(x), \boldsymbol{f}(x)> = \int_a^b \overline{\boldsymbol{f}(x)}\boldsymbol{f}(x)\; dx = \int_a^b |\boldsymbol{f}(x)|^2 dx$$

and $|\boldsymbol{f}(x)|^2 \geq 0$, hence $<\boldsymbol{f}(x), \boldsymbol{f}(x)> \geq 0$. In particular,

$$< f(x), f(x) > \; = \int_a^b |f(x)|^2 dx = 0 \quad \Rightarrow \quad |f(x)|^2 = 0$$

That is, $f(x) = 0$ $(a \leq x \leq b)$, conversely, if f is a zero function, then it is easy to see that $< f(x), f(x) >= 0$. ■

Complex inner product space, norm, distance

Definition [Norm, and distance]

Let V be a complex inner product space. The **norm of** $\mathbf{u}$ **and the distance between** $\mathbf{u}$ and $\mathbf{v}$ are defined as follows:

$$\|\mathbf{u}\| = < \mathbf{u}, \mathbf{u} >^{\frac{1}{2}}, \; d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$$

Example 5 Find the Euclidean inner product and the distance of vectors $\mathbf{u} = (2i, 0, 1+3i)$, $\mathbf{v} = (2-i, 0, 1+3i)$.

Solution

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= \overline{(2-i)}\,(2i) + 0 \cdot 0 + \overline{(1+3i)}(1+3i) \\ &= (2+i)(2i) + 0 + (1-3i)(1+3i) \\ &= 4i + 2i^2 + 1 - 9i^2 = 8 + 4i.\end{aligned}$$

$$\begin{aligned}d(\mathbf{u}, \mathbf{v}) &= \|\mathbf{u} - \mathbf{v}\| \\ &= \sqrt{|2i - (2-i)|^2 + |0-0|^2 + |(1+3i) - (1+3i)|^2} \\ &= \sqrt{|-2+3i|^2 + 0 + 0} \\ &= \sqrt{4+9} = \sqrt{13}.\end{aligned}$$ ■

Example 6 From Example 4, we let $[a, b] = [0, \pi]$ and $f(x) = e^{imx}$. Find the norm of $f(x)$.

Solution

$$\|f\| = \langle f(x), f(x)\rangle^{\frac{1}{2}} = \left(\int_0^{\pi} \overline{e^{imx}} e^{imx} dx\right)^{\frac{1}{2}}$$
$$= \left(\int_0^{\pi} e^{-imx} e^{imx} dx\right)^{\frac{1}{2}} = \left(\int_0^{\pi} dx\right)^{\frac{1}{2}} = \sqrt{\pi}\,. \quad ■$$

Cauchy-Schwarz inequality and the triangle inequality

Theorem 9.2.1

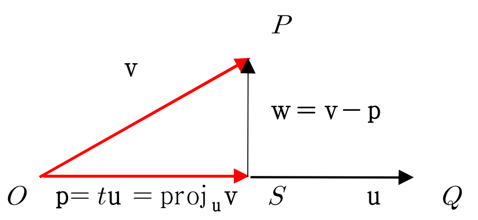
Let V be a complex inner product space. For any u, v in V, the following hold.

(1) $|\langle \mathrm{u}, \mathrm{v}\rangle| \le \|\mathrm{u}\| \|\mathrm{v}\|$. (Cauchy-Schwarz inequality)

(2) $\|\mathrm{u}+\mathrm{v}\| \le \|\mathrm{u}\| + \|\mathrm{v}\|$. (triangle inequality)

Proof We prove (1) only and leave the proof of (2) as an exercise.

If $\mathrm{u}=0$, $|\langle \mathrm{u}, \mathrm{v}\rangle| = 0 = \|\mathrm{u}\| \|\mathrm{v}\|$. Hence (1) holds. Let $\mathrm{u} \neq \mathbf{0}$ and $\mathrm{p} = \mathrm{proj}_{\langle \mathrm{u}\rangle}\mathrm{v}$, $\mathrm{w} = \mathrm{v} - \mathrm{p}$. Then $\langle \mathrm{w}, \mathrm{p}\rangle = 0$ and $\mathrm{p} = \dfrac{\langle \mathrm{v}, \mathrm{u}\rangle}{\|\mathrm{u}\|^2}\mathrm{u}$. Thus we have the following.



$$0 \le \langle \mathrm{w}, \mathrm{w}\rangle = \langle \mathrm{w}, \mathrm{v}-\mathrm{p}\rangle = \langle \mathrm{w}, \mathrm{v}\rangle - \langle \mathrm{w}, \mathrm{p}\rangle$$
$$= \langle \mathrm{w}, \mathrm{v}\rangle = \langle \mathrm{v}-\mathrm{p}, \mathrm{v}\rangle = \langle \mathrm{v}, \mathrm{v}\rangle - \langle \mathrm{p}, \mathrm{v}\rangle$$
$$= \|\mathrm{v}\|^2 - t\langle \mathrm{u}, \mathrm{v}\rangle = \|\mathrm{v}\|^2 - \frac{\langle \mathrm{v}, \mathrm{u}\rangle}{\|\mathrm{u}\|^2}\langle \mathrm{u}, \mathrm{v}\rangle.$$

Thus, as $\|\mathrm{v}\|^2\|\mathrm{u}\|^2 \ge \overline{\langle \mathrm{u}, \mathrm{v}\rangle}\langle \mathrm{u}, \mathrm{v}\rangle = |\langle \mathrm{u}, \mathrm{v}\rangle|^2$, (1) holds. ■

Example 7

Let $\mathrm{u}=(1+i,\ 0,\ 2-i)$, $\mathrm{v}=(2,\ 3i,\ i)$ be vectors in C^2. Answer the following.

(1) Compute the Euclidean inner product $<\mathrm{u},\ \mathrm{v}>$, $\|\mathrm{u}\|, \|\mathrm{v}\|, \|\mathrm{u}+\mathrm{v}\|$.

(2) Show that u and v are linearly independent.

Solution

(1) $<\mathrm{u},\ \mathrm{v}>=(2)(1+i\ )+0+(\bar{i}\ )(2-i\ )=2+2i-2i+i^2=1$

$$\|\mathrm{u}\|=\sqrt{|1+i\,|^2+|2-i\,|^2}=\sqrt{2+5}=\sqrt{7}$$

$$\|\mathrm{v}\|=\sqrt{|2|^2+|3i\,|^2+|i\,|^2}=\sqrt{4+9+1}=\sqrt{14}$$

$$\|\mathrm{u}+\mathrm{v}\|=\|(3+i,3i,\ 2)\|=\sqrt{|3+i\,|^2+|3i\,|^2+|2|^2}$$

$$=\sqrt{3^2+1^2+3^2+2^2}=\sqrt{23}$$

(2) If $\alpha\mathrm{u}+\beta\mathrm{v}=\mathbf{0}$ for any scalar $\alpha,\ \beta\in C$, then

$$\alpha(1+i,\ 0,\ 2-i)+\beta(2,\ 3i,\ i)=(0,\ 0,\ 0)$$

$$\Rightarrow\ (\alpha+2\beta)+\alpha i=0,\ 3\beta i=0,\ 2\alpha+(\beta-\alpha)i=0.$$

So $\alpha=0$, $\beta=0$. Thus u and v are linearly independent. ■

Example 8

Let $\mathrm{u}=(1+i,\ 0,\ 2-i\ )$, $\mathrm{v}=(2,\ 3i,\ i\ )$ be vectors in C^2. Check that the Cauchy-Schwarz inequality and the triangle inequality hold.

Solution

Since $|<\mathrm{u},\ \mathrm{v}>|=1$ and $\|\mathrm{u}\|\,\|\mathrm{v}\|=\sqrt{98}>1$, the Cauchy-Schwarz inequality holds.

Also since $\|\mathrm{u}+\mathrm{v}\|=\sqrt{23}\le\sqrt{7}+\sqrt{14}=\|\mathrm{u}\|+\|\mathrm{v}\|$, the triangle inequality holds. ■

Example 9

[Cauchy-Schwarz inequality in C^n and $\mathcal{E}([a,\ b],\ C)$]

(1) Let C^n be a complex inner product space with the Euclidean inner product. Let $\mathrm{u}=(a_1,\ \ldots,\ a_n)$, $\mathrm{v}=(b_1,\ \ldots,\ b_n)$ be in C^n. Then

$$|<\mathrm{u},\ \mathrm{v}>| = \left|\sum_{i=1}^{n} a_i \overline{b_i}\right| \leq \left(\sum_{i=1}^{n} |a_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} |b_i|^2\right)^{\frac{1}{2}} = \|\mathrm{u}\|\,\|\mathrm{v}\|$$

Hence the Cauchy-Schwarz inequality holds. ■

(2) Let $\mathrm{u} = f(x),\ \mathrm{v} = g(x) \in \mathcal{E}([a,\ b],\ C)$. As in Example 4, since the inner product is given by

$$|<\mathrm{u},\ \mathrm{v}>| = \left|\int_a^b \overline{g(x)} f(x)\, dx\right| \leq \left(\int_a^b |f(x)|^2 dx\right)^{\frac{1}{2}} \left(\int_a^b |g(x)|^2 dx\right)^{\frac{1}{2}}$$
$$= \|\mathrm{u}\|\|\mathrm{v}\|$$

the Cauchy-Schwarz inequality holds. ■

Example 10

[triangle inequality] Consider the inner products given in Example 3 and Example 4.

(1) Let $\mathrm{u} = (a_1,\ \ldots,\ a_n),\ \mathrm{v} = (b_1,\ \ldots,\ b_n) \in C^n$. Then the triangle inequality holds. That is,

$$\|\mathrm{u}+\mathrm{v}\| = \left(\sum_{i=1}^{n} |a_i + b_i|^2\right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^{n} |a_i|^2\right)^{\frac{1}{2}} + \left(\sum_{i=1}^{n} |b_i|^2\right)^{\frac{1}{2}}$$
$$= \|\mathrm{u}\| + \|\mathrm{v}\|.$$ ■

(2) Let $\mathrm{u} = f(x),\ \mathrm{v} = g(x) \in \mathcal{E}([a,\ b],\ C)$. Then the triangle inequality holds. That is,

$$\|\mathrm{u}+\mathrm{v}\| = \left(\int_a^b |f(x) + g(x)|^2\, dx\right)^{\frac{1}{2}}$$
$$\leq \left(\int_a^b |f(x)|^2 dx\right)^{\frac{1}{2}} + \left(\int_a^b |g(x)|^2 dx\right)^{\frac{1}{2}}$$
$$= \|\mathrm{u}\| + \|\mathrm{v}\|.$$ ■

9.3 Isomorphism

- Reference site: http://youtu.be/frOcceYb2fc, http://youtu.be/Y2lhCID0XS8
- Lab site: http://matrix.skku.ac.kr/knou-knowls/cla-week-14-sec-9-3.html

We generalize the definition of a linear transformation on $\mathbb{R}^n$ to a general vector space V. A special attention will be given to both injective and surjective linear transformations.

Definition

Let V and W be vector spaces over $\mathbb{R}$. $T: V \to W$ be a map from a vector space V to a vector space W. If T satisfies the following conditions, it is called a linear transformation.

(1) $T(c\mathbf{u}) = cT(\mathbf{u})$ for every $\mathbf{u}$ in V and c in $\mathbb{R}$.
(2) $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for every $\mathbf{u}, \mathbf{v}$ in V.

- If $V = W$, then the linear transformation T is called a linear operator.

Theorem 9.3.1

If $T: V \to W$ is a linear transformation, Then we have the following:

(1) $T(\mathbf{0}) = \mathbf{0}$.
(2) $T(-\mathbf{u}) = -T(\mathbf{u})$.
(3) $T(\mathbf{u}-\mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$.

Example 1 If $T: V \to W$ satisfies that $T(\mathbf{v}) = \mathbf{0}$ for any $\mathbf{v} \in V$, then it is a linear transformation, called the zero transformation. Also, if $T: V \to V$ satisfies that $T(\mathbf{v}) = \mathbf{v}$ for any $\mathbf{v} \in V$, then it is a linear transformation, called the identity operator. ■

Example 2 Define $T : V \to V$ by $T(\mathrm{v}) = k\mathrm{v}$ (k a scalar). Then T is a linear transformation. The following two properties hold.

(1) $T(c\mathrm{u}) = k(c\mathrm{u}) = c(k\mathrm{u}) = cT(\mathrm{u})$
(2) $T(\mathrm{u}+\mathrm{v}) = k(\mathrm{u}+\mathrm{v}) = k\mathrm{u}+k\mathrm{v} = T(\mathrm{u})+T(\mathrm{v})$

If $0 < k < 1$, then T is called a contraction and if $k > 1$, then it is called a dilation. ■

Example 3 Let $\mathcal{E}(\mathbb{R})$ be the vector space of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$ and V be the subspace of $\mathcal{E}(\mathbb{R})$ consisting of differentiable functions. Define $D : V \to V$ by $D(f) = f'$. Then D is a linear transformation and called a derivative operator. ■

Example 4 Let V the subspace of $\mathcal{E}(\mathbb{R})$ consisting of differentiable functions. Define $J : V \to W$ by $J(f) = \int_0^x f(t)\,dt$. Then J is linear transformation. ■

Kernel and Range

Definition [Kernel and Range]

Let $T : V \to W$. Define

$$\ker T = \{\mathrm{v} \in V \mid T(\mathrm{v}) = 0\},\ \operatorname{Im} T = \{T(\mathrm{v}) \in W \mid \mathrm{v} \in V\}$$

$\ker T$ is called the **kernel** and $\operatorname{Im} T$ the **range**.

Example 5 If $T : V \to W$ is the zero transformation, $\ker T = V$ and $\operatorname{Im} T = \{0\}$. ■

Example 6

If $T: V \to V$ is the identity operator, $\ker T = \{0\}$ and $\operatorname{Im} T = V$. ■

Example 7

Let D be the derivative operator defined by $D(f) = f'$ as in Example 3. $\ker D$ is "the set of all constant functions defined on $(-\infty, \infty)$" and $\operatorname{Im} T$ is "the set of all continuous functions, that is, $Cont(-\infty, \infty)$". ■

Basic properties of kernel and range

Theorem 9.3.2

If $T: V \to W$ is a linear transformation, $\ker T$ and $\operatorname{Im} T$ are subspaces of V and W, respectively.

Theorem 9.3.3

If $T: V \to W$ is a linear transformation, the following statements are equivalent.

(1) T is an injective (or one-to-one) function.
(2) $\ker T = \{0\}$.

Isomorphism

Definition

If a linear transformation $T: V \to W$ is one-to-one and onto, then it is called an isomorphism. In this case, we say that V is isomorphic to W, denoted by $V \cong W$.

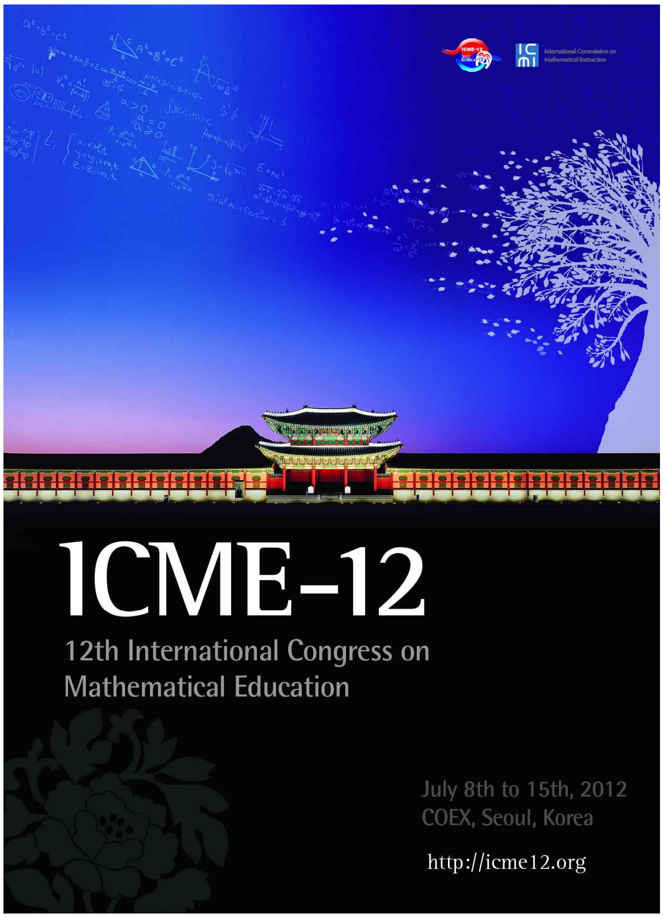
Theorem 9.3.4

Any n-dimensional real vector space is isomorphic to $\mathbb{R}^n$.

- Any n-dimensional real vector space (defined over the real set $\mathbb{R}$) is isomorphic to $\mathbb{R}^n$ and any n-dimensional complex vector space (defined over the complex set C) is isomorphic to C^n.

Example 8 We immediately obtain the following result from the above theorem.

$$P_{n-1} \cong \mathbb{R}^n, \quad M_{m \times n} \cong \mathbb{R}^{m \times n}$$ ■



[the 12th International Congress on Mathematical Education]
http://www.icme12.org/
http://matrix.skku.ac.kr/2012-Album/ICME-HPM.html

Chapter 9 Exercises

- http://matrix.skku.ac.kr/LA-Lab/index.htm
- http://matrix.skku.ac.kr/knou-knowls/cla-sage-reference.htm

Problem 1 When we define the addition and the scalar multiple on $\mathbb{R}^2$ and M_2 as follows. Check if $\mathbb{R}^2$ and M_2 are vector spaces.

(1) $(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2)$, $k(a_1, a_2) = (a_1, 0)$.

(2) $(a_1, a_2) + (b_1, b_2) = (a_1 - b_1, a_2 - b_2)$, $k(a_1, a_2) = (ka_1, ka_2)$.

(3) $\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} + \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = \begin{bmatrix} a_1 + b_1 & a_2 b_2 \\ a_3 b_3 & a_4 + b_4 \end{bmatrix}$, $k\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} = \begin{bmatrix} ka_1 & ka_2 \\ ka_3 & ka_4 \end{bmatrix}$.

(4) $\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} + \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = \begin{bmatrix} a_1 + b_1 & a_2 + b_2 \\ a_3 + b_3 & a_4 + b_4 \end{bmatrix}$, $k\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} = \begin{bmatrix} ka_1 & ka_2 \\ ka_3 & ka_4 \end{bmatrix}$.

Problem 2 Which one is a subspace of M_2?

(1) $\left\{ \begin{bmatrix} 0 & -a \\ a & 0 \end{bmatrix} \middle| a \in \mathbb{R} \right\}$.

(2) $\left\{ \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \middle| a, b \in \mathbb{R} \right\}$.

(3) $\left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| a + d = 0 \right\}$.

(4) $\left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| a + d = 1 \right\}$.

Problem 3 Let $p(t) = 2t^2 + 3t + 1$ be a vector in P_2. Write $p(t)$ as a linear combination of $p_1(t) = t^2 + t - 1$, $p_2(t) = t^2 + t + 1$, $p_3(t) = t^2 + 2t + 4$.

Solution Let $2t^2 + 3t + 1 = c_1(t^2 + t - 1) + c_2(t^2 + t + 1) + c_3(t^2 + 2t + 4)$

$= (c_1 + c_2 + c_3)t^2 + (c_1 + c_2 + 2c_3)t + (-c_1 + c_2 + 4c_3)$, $c_1, c_2, c_3 \in \mathbb{R}$.

$$\Rightarrow \begin{cases} c_1+c_2+c_3=2 \\ c_1+c_2+2c_3=3 \\ -c_1+c_2+4c_3=1 \end{cases} \Rightarrow \quad c_1=2, c_2=-1, c_3=1$$

Therefore $p(t)=2p_1(t)-p_2(t)+p_3(t)$. ■

Problem 4 Determine if the below vectors in a given vector space are linearly independent or linearly dependent.

(1) $\mathbb{R}^3$: $\mathrm{x}_1=(12,\ -12,\ 43,\ 1)$, $\mathrm{x}_2=(0,\ 12,\ -3,\ -2)$, $\mathrm{x}_3=(12,\ -12,\ 5,\ 0)$, $\mathrm{x}_4=(-2,\ 3,\ 1,\ 4)$.

(2) M_2 : $\mathrm{x}_1=\begin{bmatrix}1&2\\3&1\end{bmatrix}$, $\mathrm{x}_2=\begin{bmatrix}1&-5\\-4&0\end{bmatrix}$, $\mathrm{x}_3=\begin{bmatrix}3&-1\\2&2\end{bmatrix}$.

(3) P_3 : $p_1(t)=t^3+4t^2-2t+3$, $p_2(t)=t^3+6t^2-t+4$, $p_3(t)=3t^3+8t^2-8t+7$.

Problem 5 Let C^3 be the complex inner product space with the Euclidean inner product. Let $\mathrm{u}=(i,\ i,\ i)$, $\mathrm{v}=(i,\ -i,\ i)$. Answer the following.

(1) Compute $<\mathrm{u},\ \mathrm{v}>$.

(2) Compute $\|\mathrm{u}\|$, $\|\mathrm{v}\|$, $\|\mathrm{u}+\mathrm{v}\|$.

(3) Confirm the Cauchy-Schwarz inequality.

(4) Confirm the triangle inequality.

Solution (1) $\langle \mathrm{u},\mathrm{v}\rangle=\overline{\mathrm{v}}^T\mathrm{u}=[-i\ i\ -i]\begin{bmatrix}i\\i\\i\end{bmatrix}=(-i)i+i\times i+(-i)i=1$

(2) $\|\mathrm{u}\|=\sqrt{|i|^2+|i|^2+|i|^2}=\sqrt{1+1+1}=\sqrt{3}$

$\|\mathrm{v}\|=\sqrt{|i|^2+|-i|^2+|i|^2}=\sqrt{1+1+1}=\sqrt{3}$

$\|\mathrm{u}+\mathrm{v}\|=\|(2i,0,2i)\|=\sqrt{|2i|^2+0+|2i|^2}=\sqrt{4+0+4}=2\sqrt{2}$

(3) $\langle \mathrm{u},\mathrm{v}\rangle=1$, $\|\mathrm{u}\|\ \|\mathrm{v}\|=3$ implies $|\langle \mathrm{u},\mathrm{v}\rangle|\le\|\mathrm{u}\|\ \|\mathrm{v}\|$.

(4) $\|\mathrm{u}\|=3$, $\|\mathrm{v}\|=3$, $\|\mathrm{u}+\mathrm{v}\|=2\sqrt{2}$ => $2\sqrt{2}\le 3+3$

Triangle inequality $\|\mathrm{u}+\mathrm{v}\|\le\|\mathrm{u}\|+\|\mathrm{v}\|$ holds. ■

Problem 6 Define an inner product on $\mathbb{R}^2$ as $<\mathrm{u},\mathrm{v}>=6u_1v_1-2u_2v_1-2u_1v_2+3u_2v_2$. Compute the following. (here $\mathrm{u}=<u_1, u_2>$, $\mathrm{v}=(v_1, v_2)$)

(1) The 2×2 symmetric matrix A such that $<\mathrm{u},\mathrm{v}>=\mathrm{v}^T A\mathrm{u}$

(2) The norm $\|\mathrm{u}\|$ of $\mathrm{u}=(1, -1)$.

(3) The norm $\|\mathrm{v}\|$ of $\mathrm{v}=(4, 3)$.

(4) θ such that $\dfrac{<\mathrm{u},\mathrm{v}>}{\|\mathrm{u}\|\|\mathrm{v}\|}=\cos\theta$.

Solution (1) Let $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\langle \mathrm{u},\mathrm{v}\rangle=6u_1v_1-2u_2v_1-2u_1v_2+3u_2v_2$.

=> $\langle \mathrm{u},\mathrm{v}\rangle=\mathrm{v}^T A\mathrm{u}=[v_1\ v_2]\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}=av_1u_1+bv_1u_2+cv_2u_1+dv_2u_2$.

$\therefore\ a=6, b=-2, c=-2, d=3,\ A=\begin{bmatrix} 6 & -2 \\ -2 & 3 \end{bmatrix}$

(2) and (3) $\|\mathrm{u}\|=\sqrt{\langle \mathrm{u},\mathrm{u}\rangle}=\sqrt{[1 -1]\begin{bmatrix} 6 & -2 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}}=\sqrt{7}$

$\|\mathrm{v}\|=\sqrt{\langle \mathrm{v},\mathrm{v}\rangle}=\sqrt{[4\ 3]\begin{bmatrix} 6 & -2 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \end{bmatrix}}=\sqrt{75}$

(4) $\cos\theta=\dfrac{\langle \mathrm{u},\mathrm{v}\rangle}{\|\mathrm{u}\|\ \|\mathrm{v}\|}=\dfrac{[4\ 3]\begin{bmatrix} 6 & -2 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}}{\sqrt{7}\times\sqrt{75}}=\dfrac{17}{5\sqrt{21}}$

=> $\theta=\cos^{-1}\dfrac{17}{5\sqrt{21}}\approx 42.103^\circ$ ■

Problem 7 Tell which one is a linear transformation or not. If not, give a reason.

(1) $T: P_3\to P_4,\ T(p(x))=xp(x)$.

(2) $T: M_n\to M_n,\ T(A)=A^T$.

(3) $T: \mathcal{E}[a, b]\to\mathbb{R},\ T(f(x))=\int_a^b f(x)\,dx$.

(4) $T: M_n\to\mathbb{R},\ T(A)=\mathrm{tr}(A)$.

(5) $T : P_3 \to P_6$, $T(p(x)) = p(x)^2$.

(6) $T : V \to \mathbb{R}$, $T(\mathbf{x}) = \|\mathbf{x}\|$.

Problem 8 Find the kernel and the range of the following linear transformations.

(1) $T : P_2 \to P_3$, $T(p(x)) = x\,p(x)$.

(2) $T : \mathcal{E}[0, 1] \to R$, $T(f(x)) = \int_0^1 f(x)\,dx$.

P1 If W_1, W_2 are subspaces of a vector space V, prove that $W_1 \cap W_2$ is a subspace of V.

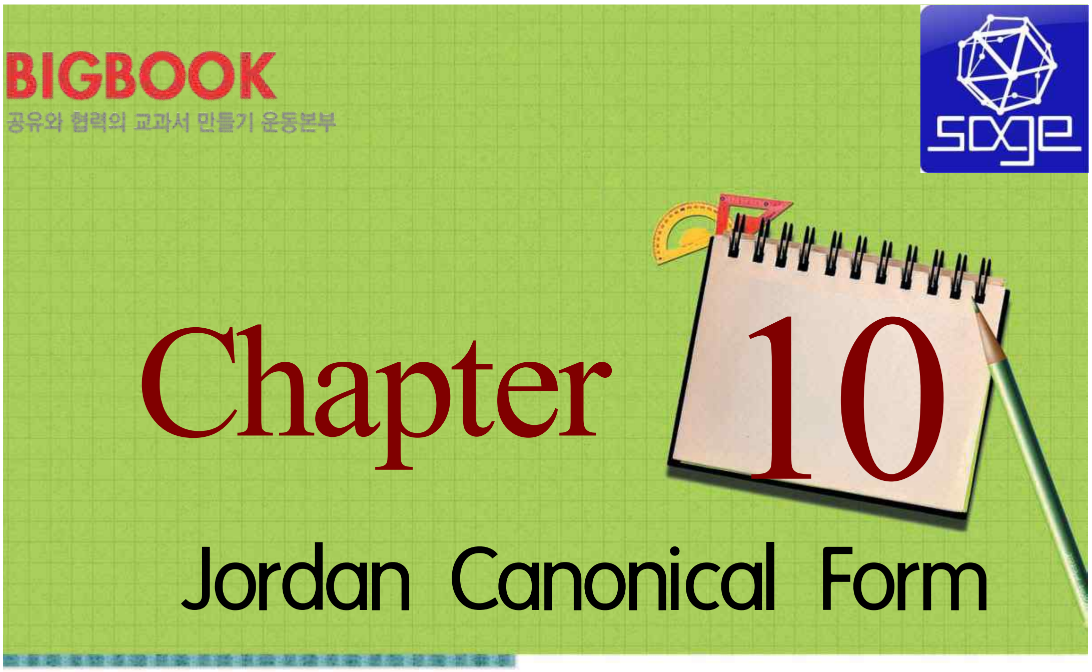
P2 Let $\mathbf{a}$ be a fixed vector and W be the set of all vectors orthogonal to $\mathbf{a}$, that is, $W = \{\mathbf{x} \in R^3 \mid \mathbf{a} \cdot \mathbf{x} = 0\}$. Show that W is a subspace of $\mathbb{R}^3$.

P3 Let C^3 be the vector space with the Euclidean inner product. Transform $\mathbf{u}_1 = (i, i, i)$, $\mathbf{u}_2 = (0, i, i)$, $\mathbf{u}_3 = (0, 0, i)$ into an orthonormal basis by using the Gram-Schmidt process.

P4 Find the standard matrix corresponding to the given linear transformation $T : M_2 \to M_2$ defined by

$$T\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{bmatrix} 2a_{21} & a_{11} + a_{21} \\ a_{12} - 2a_{21} & a_{22} \end{bmatrix}.$$

P5 Define $T : M_2(\mathbb{R}) \to M_2(\mathbb{R})$ by $T(A) = A + A^T$. Show that it is a linear transformation and find the kernel and the range of T.



If a matrix is diagonalizable, every thing is much easier. But most of matrices are not diagonalizable. The Jordan canonical form is an upper triangular matrix of a particular form called a Jordan matrix (a simple block diagonal matrix) representing an operator with respect to some basis. The diagonal entries of the normal form are the eigenvalues of the operator, with the number of times each one occurs given by its algebraic multiplicity.

Any square matrix has a Jordan normal form if the field of coefficients is extended to one containing all the eigenvalues of the matrix. Since each matrix has a corresponding Jordan canonical form which is similar to it, all computations can be done with this simple upper triangular matrix. The Jordan normal form is named after Camille Jordan, a French mathematician renowned for his work in various branches of mathematics. In this chapter, we will study how to find a Jordan matrix which is similar to any given matrix and how to find generalized eigenvectors.

10.1 Finding the Jordan Canonical Form with a Dot Diagram

- Reference video: http://youtu.be/NBLZPcWRHYI, http://youtu.be/NBLZPcWRHYI
- Practice site: http://matrix.skku.ac.kr/knou-knowls/cla-week-15-sec-10-1.html
  http://matrix.skku.ac.kr/JCF/

If a given matrix is diagonalizable, most computational problems involving that matrix and desired conclusions can be easily obtained. However, not every matrix is diagonalizable. In this section, we will introduce a method for finding the Jordan Canonical Form of a non-diagonzaliable matrix by a similarity transformation.

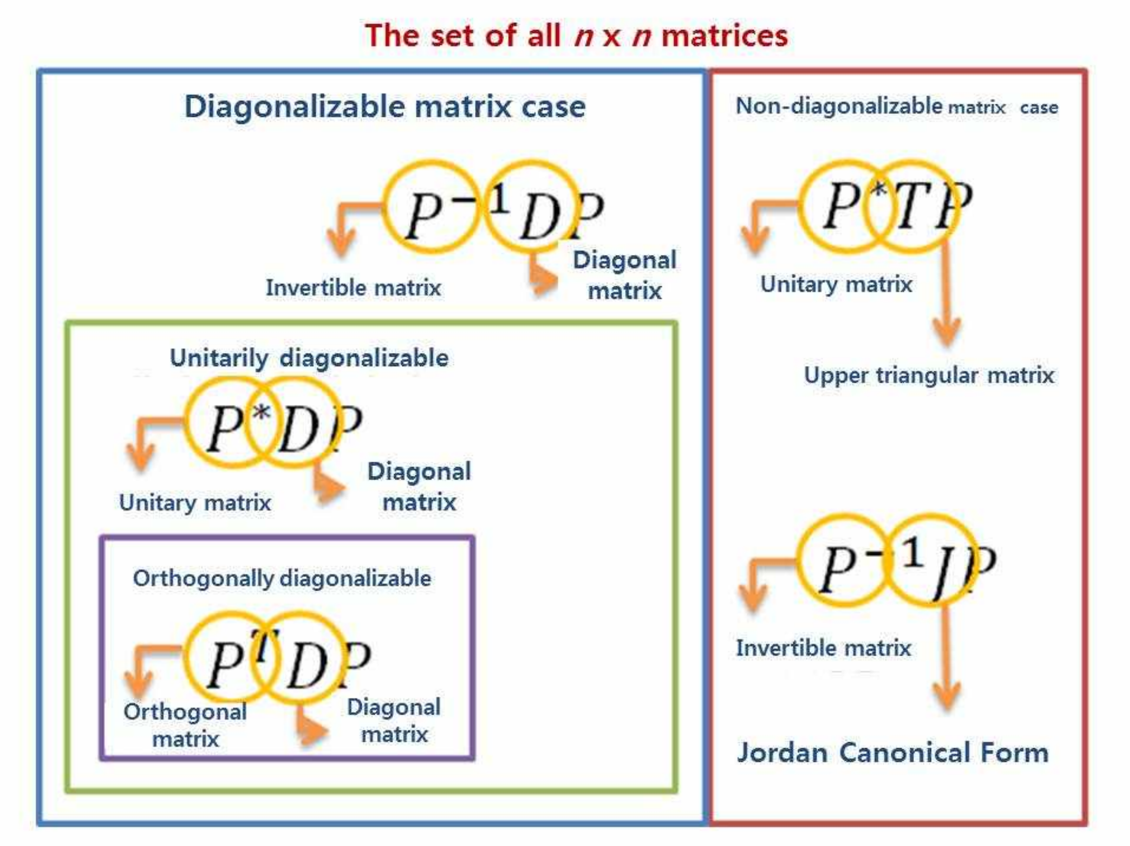
Let us review a few concepts of matrix diagonalization.

Diagonalization of a Square Matrix (*Review*)

1. Let A be an $n \times n$ matrix. Then, A is diagonalizable if and only if it has n linearly independent eigenvectors. However not all matrices are diagonalizable.
2. A normal matrix A $(AA^* = A^*A)$ is unitarily diagonalizable (that is, unitarily similar to a diagonal matrix). However, not all diagonalizable matrices are normal.
3. If a matrix A is diagonalizable, each eigenvalue of A generates an eigenspace with dimension equal to the algebraic multiplicity of that eigenvalue.

- For *every* square matrix A (not necessarily diagonalizable), one can obtain a block-diagonal matrix called the Jordan canonical form matrix that is similar to A.

For example, the matrices $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 6 & 3 & -8 \\ 0 & -2 & 0 \\ 1 & 0 & -3 \end{bmatrix}$ are non diagonalizable.



Theorem 10.1.1

Let A be an $n \times n$ matrix with t $(1 \leq t \leq n)$ linearly independent eigenvectors. Then, A is similar to a matrix

$$J_A = \begin{bmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_t \end{bmatrix}_{n \times n}$$

where $U^* A U = J_A$ for some unitary matrix U. Furthermore, we have

$$J_k = \begin{bmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{bmatrix}_{n_k \times n_k}, \quad (n_1 + n_2 + \cdots + n_t = n, \ 1 \leq k \leq t)$$

where each J_k, called a Jordan block, corresponds to an eigenvalue λ_i of A. The block diagonal matrix J_A is called the **Jordan canonical form** of A and each J_k are called Jordan blocks of J_A.

- The Jordan Canonical Form (JCF) of a matrix A is a block diagonal matrix composed of Jordan blocks, each with eigenvalues of A on its respective diagonal, 1's on its superdiagonal, and 0's elsewhere.

Remark **Properties of Jordan blocks**

1. For a given eigenvalue λ_i of an $n \times n$ matrix A, its **geometric multiplicity** is the number of linearly independent eigenvectors associated with λ_i: hence, it is the number of Jordan blocks corresponding to λ_i.

2. The sum of the sizes (i.e. orders) of all Jordan blocks corresponding to an eigenvalue λ_i is its **algebraic multiplicity**.

3. If the geometric multiplicity and algebraic multiplicity of every eigenvalue of A are equal, then the size of every Jordan block is 1×1, and

$$\text{sum of algebraic multiplicities} = \text{sum of geometric multiplicities} = \text{size of } A$$

In this case, the matrix A is diagonalizable. (This type of matrix is called a **simple matrix**.)

Example 1 The matrix J_A

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

has characteristic polynomial $\det(A - \lambda I) = (\lambda - 2)^4 (\lambda - 3)^2 \lambda^2$ and is the Jordan Form of some 8×8 square matrix A.

Notice how the algebraic multiplicities of each eigenvalue determines the number of times that eigenvalue appears along the principal diagonal of J_A : 2 appears four times, while 3 and 0 appear two times. Hence the algebraic multiplicity of 2 is 4.

You can easily verify that the sum of the sizes of all Jordan blocks corresponding to a single eigenvalue is also equal to its algebraic multiplicity.

Also, note that the geometric multiplicities of each eigenvalue determine the number of Jordan blocks corresponding to that eigenvalue.
Thus geometric multiplicities of 2, 3 and 0 are 2, 1, and 1 respectively.■

Example 2

For a 5×5 square matrix A, if A has only one eigenvalue λ with one associated linearly independent eigenvector, the Jordan form of A is the following:

$$J_A = \begin{bmatrix} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \lambda & 1 & 0 \\ 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{bmatrix}$$

This is due to the fact that the number of linearly independent eigenvectors of A determines the number of Jordan blocks in the Jordan form of A. Thus in this case the geometric multiplicity of the eigenvalue 5 is 1 where as algebraic multiplicity is 5. ■

How to find the Jordan Canonical Form

- Suppose for some matrix $A \in M_n$ with k distinct eigenvalues λ_1, λ_2, ..., λ_k, the Jordan canonical form of A, J_A, is the following:

$$J_A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & A_k \end{bmatrix}.$$

- Here, each A_i corresponds to a Jordan block with the eigenvector λ_i along its diagonal. These are called block submatrices of J_A. Now, for each eigenvalue λ_i, we have a block submatrix

$$A_i = \begin{bmatrix} J_{i,p_1} & 0 & \cdots & 0 \\ 0 & J_{i,p_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & J_{i,p_{l_i}} \end{bmatrix}$$

and, knowing its structure, we can easily find the Jordan Canonical Form J_A. The Jordan form is uniquely determined up to the order of the blocks; that is, the number and size of the Jordan blocks associated with each eigenvalue is uniquely determined, but the blocks can appear in any order along the main diagonal.

- For each eigenvalue $\lambda_i\,(i=1,2,\,\ldots,\,k)$, A_i consists of $l_i(1 \le i \le k)$ Jordan blocks; let us find the size of each $J_{i,p_t} \in M_{p_t}$, namely p_t $(1 \le t \le l_i)$. For the set of linearly independent eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{l_i}$ corresponding to λ_i, for ease of notation, let us first consider only one eigenvalue. Therefore, we let λ_i be λ and l_i be l.

- The number of Jordan blocks in A_i, l, and their corresponding sizes $p_1, p_2, \ldots, p_l$ is determined by calculating the rank of $A-\lambda I$. Without loss of generality, we take $p_1 \ge p_2 \ge \cdots \ge p_l$. Now, for the eigenvalue λ and the dimension of its corresponding λ-eigenspace (its geometric multiplicity), using l and p_t, we introduce a sequence of points to easily calculate A_i; this is called the dot diagram. The dots in the dot diagram are configured according to the following rules:

  * **Dot Diagram Properties**

1. The dot diagram consists of l columns.
2. Counting from left to right, the jth column consists of the k_j dots that correspond to the eigenvectors of x_j, starting with the initial vector at the top and continuing down to the end vector.

- Thus, the following is the dot diagram of A_i:

$$\begin{array}{lllll}
\bullet\ (A-\lambda I)^{k_{p_1}-1}(\mathbf{x}_1) & \bullet\ (A-\lambda I)^{k_{p_2}-1}(\mathbf{x}_2) & \cdots & \bullet\ (A-\lambda I)^{k_{p_l}-1}(\mathbf{x}_l) \\
\bullet\ (A-\lambda I)^{k_{p_1}-2}(\mathbf{x}_1) & \bullet\ (A-\lambda I)^{k_{p_2}-2}(\mathbf{x}_2) & \cdots & \vdots \\
\quad\vdots & \quad\vdots & & \bullet\ \mathbf{x}_l
\end{array}$$

- $(A-\lambda I)(\mathbf{x}_1)$ - $\mathbf{x}_2$
- $\mathbf{x}_1$

- Here, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_l$ are the linearly independent eigenvectors associated to the eigenvalue λ_i. Let r_j denote t the number of dots in the jth row of the dot diagram; then, r_1 is the number of Jordan blocks of size *at least* 1×1 , r_2 is the number of Jordan blocks of size *at least* 2×2, and r_{p_1} is the number of Jordan blocks of size *at least* $p_1\times p_1$. Thus, $r_1 \geq r_2 \geq \cdots \geq r_{p_1}$. Refer to Theorem 10.1.2 and 10.1.3, and consider the example below.

Example 3 For a 9×9 matrix A_i, the number of Jordan blocks contained in A_i is l and the size of the Jordan blocks is completely determined by $p_1, p_2, \ldots, p_l$. To see this, take $l=4$ and $p_1=3, p_2=3, p_3=2, p_4=1$. Then, following the sequence of block sizes,

$$A_i = \begin{bmatrix} \lambda_i & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_i & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_i & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_i & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_i & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_i & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_i & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_i \end{bmatrix}$$

is uniquely determined.

To find the dot diagram of A_i, since $l=4, p_1=3, p_2=3, p_3=2$ and $p_4=1$, the dot diagram of is:

• • • • (Number of Jordan blocks: 4)
• • •
• • ■

Theorem 10.1.2

The number of dots in the first r rows of the dot diagram for λ_i is equal to the dimension of solution space of $(A-\lambda_i I)^r \mathbf{x} = \mathbf{0}$ (i.e. the nullity of $(A-\lambda_i I)^r$).

- nullity$(A-\lambda_i I)^r$ = nullity$(J_A-\lambda_i I)^r$

Theorem 10.1.3

For $A \in M_n(C)$, let r_j denote the number of dots in the jth row of the dot diagram of λ_i. Then, the following are true.

(1) $r_1 = n - \text{rank}(A-\lambda_i I)$.

(2) If $j > 1$, $r_j = \text{rank}((A-\lambda_i I)^{j-1}) - \text{rank}((A-\lambda_i I)^j)$.

Proof By **Theorem 10.1.2**,

$$r_1 + r_2 + \cdots + r_j = \text{nullity}\ ((A-\lambda_i I)^j) = n - \text{rank}((A-\lambda_i I)^j)\ \text{(provided } j \geq 1)$$

Also, $r_1 = n - \text{rank}(A-\lambda_i I)$ and

$$\begin{aligned} r_j &= (r_1 + r_2 + \cdots + r_j) - (r_1 + r_2 + \cdots + r_{j-1}) \\ &= [n - \text{rank}((A-\lambda_i I)^j)] - [n - \text{rank}((A-\lambda_i I)^{j-1})] \\ &= \text{rank}((A-\lambda_i I)^{j-1}) - \text{rank}((A-\lambda_i I)^j),\ j > 1. \end{aligned}$$

(The number of dots in each row, r_j, means the number of blocks of size *at least* $j \times j$) ■

- From Theorem 10.1.3, let's see how the dot diagram for each λ_i is completely determined by the matrix A.

Example 4 Find the Jordan Canonical Form of A.

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

Solution

```
A=matrix(4, 4, [2, -1, 0, 1, 0, 3, -1, 0, 0, 1, 1, 0, 0, -1, 0, 3])
print A.charpoly().factor()
print A.eigenvalues()
```

```
(x - 3) * (x - 2)^3
[3, 2, 2, 2]
```

The matrix A has characteristic polynomial $\det(A-\lambda I)=(\lambda-3)(\lambda-2)^3$, so A has two distinct eigenvalues λ_1, λ_2.

Here $\lambda_1=3$ has algebraic multiplicity 1, and $\lambda_2=2$ has algebraic multiplicity 3. Thus, the dot diagram for λ_1 has 1 dot

$$\bullet$$

and A_1 has one 1×1 Jordan block. That is, $A_1=[3]$. As well, the dot diagram for λ_2 has 3 dots, and

```
E=identity_matrix(4)
print (A-2*E).rank()
print ((A-2*E)^2).rank()
```

```
2
1
```

$$r_1=4-\mathrm{rank}(A-2I)=4-\mathrm{rank}\begin{bmatrix}0&-1&0&1\\0&1&-1&0\\0&1&-1&0\\0&-1&0&1\end{bmatrix}=4-2=2,$$

$r_2=\mathrm{rank}(A-2I)-\mathrm{rank}\,[(A-2I)^2]=2-1=1$.

Thus, the dot diagram for λ_2 is the following.

$r_1=2$: $\bullet$ $\quad\bullet$ (number of Jordan blocks: 2)
$r_2=1$: $\bullet$

A_2 has one 2×2 Jordan block and one 1×1 Jordan block. That is,

$$A_2 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Hence, the Jordan Canonical form of A is

$$\therefore \quad J_A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \qquad \square$$

Sage http://sage.skku.edu and http://mathlab.knou.ac.kr:8080/

```
A=matrix(4, 4, [2, -1, 0, 1, 0, 3, -1, 0, 0, 1, 1, 0, 0, -1, 0, 3])
J=A.jordan_form()    # Jordan Canonical Form
print J
```

```
[3|0 0|0]
[-+---+-]
[0|2 1|0]
[0|0 2|0]
[-+---+-]
[0|0 0|2]
```
■

Example 5 Find the Jordan Canonical Form of A.

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

Solution

The matrix A has characteristic polynomial $\det(A-\lambda I) = (\lambda-2)^2(\lambda-4)^2$, so there are two distinct eigenvalues of A, $\lambda_1 = 4$와 $\lambda_2 = 2$, each with algebraic multiplicity 2. For $\lambda_1 = 4$,

$$r_1 = 4 - \operatorname{rank}(A-4I) = 4 - 3 = 1$$

Therefore, the dot diagram for λ_1 is the following.

$$r_1 = 1 \; : \; \bullet \qquad \text{(Number of Jordan blocks: 1)}$$
$$r_2 = 1 \; : \; \bullet$$

So, $A_1 = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$.

For $\lambda_2 = 2$, $r_1 = 4 - \text{rank}(A - 2I) = 4 - 2 = 2$. r_2 is 0($\because$ The number of dots is $2 = 2 + (0)$). Therefore, the dot diagram for λ_2 is the following:

$$r_1 = 2 \; : \; \bullet \; \bullet \quad \text{(Number of Jordan blocks = 2)}$$

So, $A_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

Thus, the Jordan Canonical Form of A is $J_A = \begin{bmatrix} 4 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$. □

Sage http://sage.skku.edu and http://mathlab.knou.ac.kr:8080/

```
A=matrix(4, 4, [2, -2, -2, -2, -4, 0, -2, -6, 2, 1, 3, 3, 2, 3, 3, 7])
J=A.jordan_form()    # Jordan Canonical Form
print J
```

```
[4 1|0|0]
[0 4|0|0]
[--+-+-]
[0 0|2|0]
[--+-+-]
[0 0|0|2]
```
■

[Remark] Jordan Canonical Form Learning Materials

- http://matrix.skku.ac.kr/2012-mobile/E-CLA/10-1.html
- http://matrix.skku.ac.kr/2012-mobile/E-CLA/10-1-ex.html

http://matrix.skku.ac.kr/JCF/index.htm

10.2 Jordan Canonical Form and Generalized Eigenvectors

- Reference video: http://www.youtube.com/watch?v=yJ7n0icjtNA
- Practice site: http://matrix.skku.ac.kr/knou-knowls/cla-week-15-sec-10-2.html
  http://matrix.skku.ac.kr/sglee/03-Note/GeneralizedEV-f.pdf
  http://matrix.skku.ac.kr/MT-04/chp8/3p.html

In **Section 10.1**, for any $n \times n$ matrix A, we discussed the theory and method for finding a matrix J_A, called the Jordan Canonical form, such that $P^{-1}AP = J_A$. In this section, we will examine a method for finding the matrix P in the above equation. This method utilizes the concept of generalized eigenvectors.

The following matrix was referred from the wiki http://en.wikipedia.org/wiki/Jordan_normal_form.

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

Consider the matrix A. The Jordan normal form is obtained by some similarity transformation $P^{-1}AP = J_A$, i.e. $AP = PJ_A$.

Let P have column vectors $\mathbf{p}_i$, $i = 1, \ldots, 4$, then

$$A[\mathbf{p}_1 : \mathbf{p}_2 : \mathbf{p}_3 : \mathbf{p}_4] = [\mathbf{p}_1 : \mathbf{p}_2 : \mathbf{p}_3 : \mathbf{p}_4]\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix} = [\mathbf{p}_1 : 2\mathbf{p}_2 : 4\mathbf{p}_3 : \mathbf{p}_3 + \mathbf{p}_4].$$

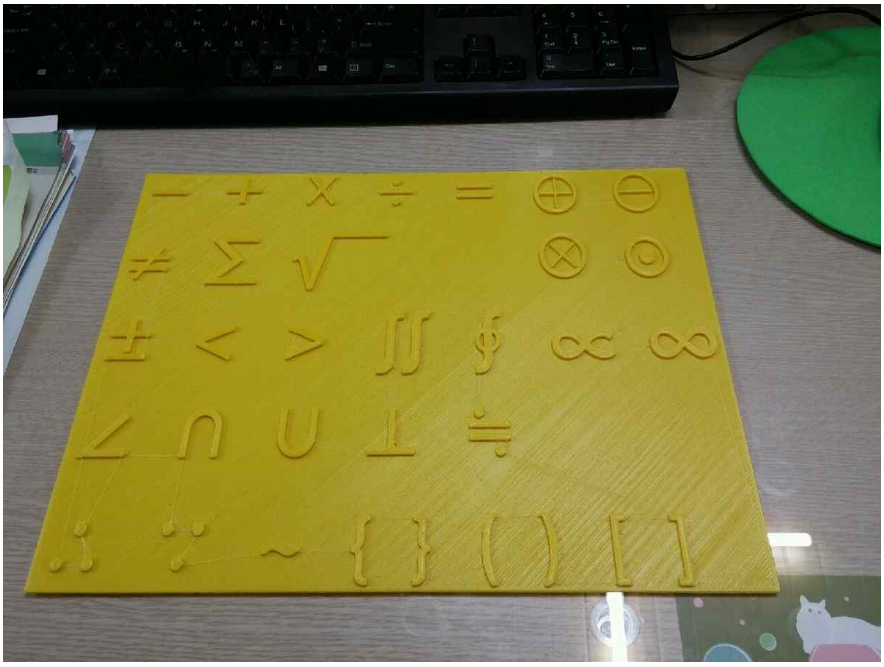
We see that

$$(A - 1I)\mathbf{p}_1 = \mathbf{0}$$
$$(A - 2I)\mathbf{p}_2 = \mathbf{0}$$
$$(A - 4I)\mathbf{p}_3 = \mathbf{0}$$
$$(A - 4I)\mathbf{p}_4 = \mathbf{p}_3.$$

For $i = 1, 2, 3$, we have $\mathbf{p}_i \in \operatorname{Ker}(A - \lambda I)$, i.e. $\mathbf{p}_i$ is an eigenvector of A corresponding to the eigenvalue λ_i. For $i = 4$, multiplying both sides by $(A - 4I)$ gives $(A - 4I)^2\mathbf{p}_4 = (A - 4I)\mathbf{p}_3$. But $(A - 4I)\mathbf{p}_3 = \mathbf{0}$, so $(A - 4I)^2\mathbf{p}_4 = \mathbf{0}$. Thus,

$\mathbf{p}_4 \in \mathrm{Ker}(A-\lambda I)^2$. Vectors like $\mathbf{p}_4$ are called **generalized eigenvectors** of A. Thus, given an eigenvalue λ, its corresponding Jordan block gives rise to a Jordan chain. The generator, or lead vector (say, $\mathbf{p}_r$) of the chain is a generalized eigenvector such that $(A-\lambda I)^r \mathbf{p}_r = \mathbf{0}$, where r is the size of the Jordan block. The vector $\mathbf{p}_1 = (A-\lambda I)^{r-1}\mathbf{p}_r$ is an eigenvector corresponding to λ. In general, $\mathbf{p}_i$ is the preimage of $\mathbf{p}_{i-1}$ under $A-\lambda I$. So the lead vector generates the chain via multiplication by $A-\lambda I$. Therefore, the statement that every square matrix A can be put in Jordan normal form is equivalent to the claim that there exists a basis consisting only of eigenvectors and generalized eigenvectors of A.

PS: More details about the Jordan Canonical Form can be found at
http://www.uio.no/studier/emner/matnat/math/MAT2440/v11/undervisningsmateriale/genvectors.pdf.



10.3 Jordan Canonical Form and CAS

- Reference video: http://youtu.be/LxY6RcNTEE0, http://youtu.be/LxY6RcNTEE0
- Practice site: http://matrix.skku.ac.kr/knou-knowls/cla-week-15-sec-10-3.html

In practice, in order to find the Jordan Canonical Form of a 10×10 matrix, you need to find the roots of a characteristic polynomial of degree 10 - the factorization and rigorous calculation of these roots is impossible. Moreover, a 10×10 matrix requires us to calculate many exponents and coefficients. In order to calculate these coefficients, the Gaussian elimination and related computations can be performed by various computer programs - e.g. HLINPRAC, MATHEMATICA, MATLAB, and the recently developed open-source program, Sage. The use of software for computationally complex mathematics is necessary in an increasingly technological society.

The following links provide more information about the Jordan Canonical Form and tools that allow you to explicitly find the Jordan Canoncial Form for a given matrix without arduous calculations by hand.

1. Theory and tools : http://matrix.skku.ac.kr/JCF/index.htm

2. Jordan Canonical Form; an algorithmic approach:
http://matrix.skku.ac.kr/JCF/JCF-algorithm.html

3. Jordan Canonical Form (step by step) tool:
http://matrix.skku.ac.kr/JCF/JordanCanonicalForm-SKKU.html

4.CAS Tool : http://matrix.skku.ac.kr/2014-Album/MC-2.html

"The man ignorant of mathematics will be increasingly limited in his grasp of the main forces of civilization."
- John George Kemeny (1926 -1992)

A Jewish-Hungarian American mathematician, computer scientist, and educator best known for co-developing **the BASIC programming language** in 1964 and pioneered the use of computers in college education.

Chapter 10 Exercises

- http://matrix.skku.ac.kr/LA-Lab/index.htm
- http://matrix.skku.ac.kr/knou-knowls/cla-sage-reference.htm

Problem 1 Let A be a 5×5 matrix with the only one eigenvalue λ with algebraic multiplicity of 5. Find all possible types of Jordan Canonical forms of A when the number of linearly independent eigenvectors corresponding λ is 2.

Problem 2 For the given Jordan Canonical form J_A, calculate the following:

$$J_A = \begin{bmatrix} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \lambda & 1 & 0 \\ 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{bmatrix}$$

(1) $J_A - \lambda I$

(2) $(J_A - \lambda I)^2$

(3) $(J_A - \lambda I)^3$

(4) $(J_A - \lambda I)^4$

[Problem 3 - 8] Find the Jordan Canonical Form of the given matrix.

Problem 3 $\begin{bmatrix} 5 & 5 & 5 & 5 & 5 \\ 5 & 5 & 5 & 5 & 5 \\ 5 & 5 & 5 & 5 & 5 \\ 5 & 5 & 5 & 5 & 5 \\ 5 & 5 & 5 & 5 & 5 \end{bmatrix}$

Solution Sage :

```
A=matrix(QQ,5,5,[5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5,5])
J=A.jordan_form()
print J
```

[25| 0| 0| 0| 0]
[0| 0| 0| 0| 0]
[0| 0| 0| 0| 0]
[0| 0| 0| 0| 0]
[0| 0| 0| 0| 0] ■

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

Problem 5 $\begin{bmatrix} 5 & -3 & -2 \\ 8 & -5 & -4 \\ -4 & 3 & 3 \end{bmatrix}$

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

Problem 7 $\begin{bmatrix} 1 & 3 & -6 & 4 & 3 \\ 0 & 1 & 2 & 8 & 2 \\ 0 & 0 & 2 & 0 & -8 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}$

Problem 8 $\begin{bmatrix} 2 & -4 & 2 & 2 \\ -2 & 0 & 1 & 3 \\ -2 & -2 & 3 & 3 \\ -2 & -6 & 3 & 7 \end{bmatrix}$

Solution Sage : --

```
A=matrix(QQ,[[2,-4,2,2],[-2,0,1,3],[-2,-2,3,3],[-2,-6,3,7]])
print A.eigenvalues()
E=identity_matrix(4);
print (A-2*E).rank(); print (A-4*E).rank();print ((A-4*E)^2).rank()
```

[4, 4, 2, 2]

2

3

2

rank$(A-4I)$=3 => $r_1 = n - \text{rank}(A-4I) = 4-3 = 1$.

$r_1 = 1$: · (Number of Jordan block : 1)

$r_2 = 1$: ·

rank$(A-2I) = 2$ => $r_1 = n -$ rank$(A-2I) = 4-2 = 2$.

$r_1 = 2$: · · (Number of Jordan block : 2)

$$\therefore J_A = \begin{bmatrix} 4 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$ ■

Problem 9 (**⊕** means **the matrix direct sum** of n matrices constructs a block diagonal matrix, **http://mathworld.wolfram.com/MatrixDirectSum.html,** from a set of square matrices.)

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

Problem 10 $$A = \begin{bmatrix} 0&-3&1&2 \\ -2&1&-1&2 \\ -2&1&-1&2 \\ -2&-3&1&4 \end{bmatrix} \oplus \begin{bmatrix} 1&3&-6&4&3 \\ 0&1&2&8&2 \\ 0&0&2&0&-8 \\ 0&0&0&1&1 \\ 0&0&0&0&4 \end{bmatrix} \oplus \begin{bmatrix} 2&-4&2&2 \\ -2&0&1&3 \\ -2&-2&3&3 \\ -2&-6&3&7 \end{bmatrix} \oplus$$

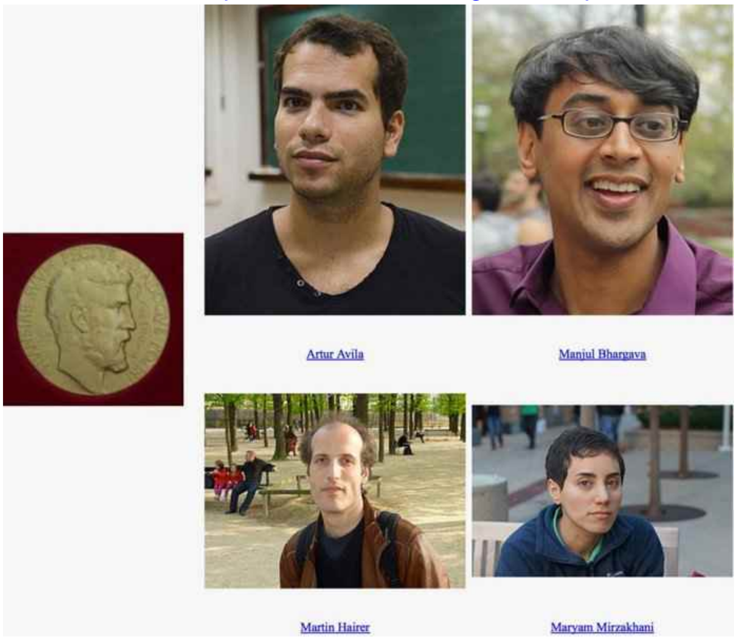
– Quotes by Great Mathematicians:

http://prezi.com/z0hgrw8a6wql/define-math/

ILAS 2014 Official Photo : http://matrix.skku.ac.kr/2014-Album/ILAS-2014/
ILAS 2014 Movie A – Registration and Presentations : http://youtu.be/asJfRFYWPrk
ILAS 2014 Movie B – Tour : http://youtu.be/bidJNagmRXQ
ILAS 2014 Movie C – Banquet : http://youtu.be/10fDqWA-vVA
ILAS 2014 Movie D – Group Photo : http://youtu.be/6IlS8U6i_8E
ILAS 2014 Movie E – Conference Preparations : http://youtu.be/UMwLCtSGByI

ICM 2014, COEX, Seoul, Fields Medalists:

http://matrix.skku.ac.kr/2014-Album/2014-ICM-SGLee/
https://www.facebook.com/SEOULICM2014
http://www.icm2014.org/en/vod/videos
http://www.icm2014.org/en/vod/public



Appendix

http://matrix.skku.ac.kr/Cal-Book/Appnd/index.htm

References

H. Anton, R. Busby, *Contemporary Linear Algebra*, Anton Textbooks Inc., 2003.

Duk-Sun Kim, Sang-Gu Lee, Greg Markowsky, Mobile Sage-Math for Linear Algebra and its Application, *Electronic Journal of Mathematics & Technology* 4 (2010), No. 3, 1-13.

Kyung-Won Kim, Sang-Gu Lee, Shaowei Sun, Modeling of Mobile Sage and Graphing Calculator, *Journal of Modern Education Review* 3 (2013), No. 12, 918-925.

Jin Ho Kwak, Sungpyo Hong, *Linear Algebra*, Birkhäuser, 1997.

Sang-Gu Lee, *Contemporary Linear algebra (3rd Edition)*, Kyungmoon Books, 2009.

Sang-Gu Lee, Kyung-Won Kim, Jae Hwa Lee, Sage matrix calculator and full Sage contents for linear algebra, *Korean J. Math.* 20 (2013), No. 4, 503-521.

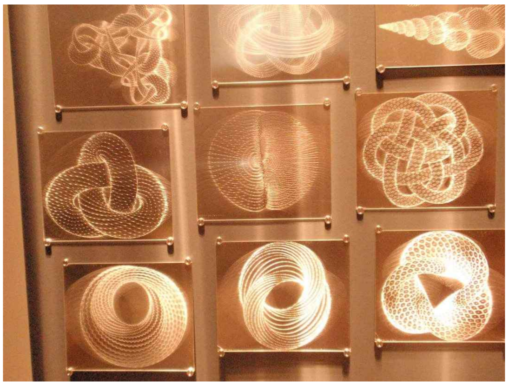
Steven J. Leon, Eugene Herman, Richard Faulkenberry, *ATLAST: Computer Exercise for Linear Algebra*, Prentice Hall Inc., 1996.

Steven J. Leon, *Linear Algebra with Applications*, Prentice Hall Inc., 2002.

Gilbert Strang, *Linear Algebra and its Application*, Thomson Learning Inc., 1988.

MOMATH MUSEUM OF MATHEMATICS

[National Museum of Mathematics, NYC] http://momath.org/



Sample Exam

- Reference video: http://youtu.be/CLxjkZuNJXw
- Practice site:
  http://matrix.skku.ac.kr/CLA-Exams-Sol.pdf,
  http://matrix.skku.ac.kr/2015-album/2015-LA-S-Exam-All-Sol.pdf
  http://matrix.skku.ac.kr/2012-album/2012-LA-Lectures.htm

* Provides basic commands if you use Sage in your test.

```
<Sage Linear Algebra partial commands list>
var('a,b,c,d')                 # define variables
eq1=3*a+3*b==12                # define equation1
eq2=5*a+2*b==13                # define equation2
solve([eq1, eq2], a,b)         # solve system of
                                 equations
A=matrix(CDF, 3, 3, [3, 0, 0, 0, 0, 2, 0, 3, 4]);
                               # define matrix
A.echelon_form()               # RREF
A.inverse()                    # inverse matrix
A.det()                        # determinant
A.adjoint()                    # adjoint matrix
A.eigenvalues()                # eigenvalues
A.eigenvectors_right()         # eigenvectors
A.charpoly()                   # characteristic equation

P,L,U=A.LU()                   # LU decomposition
(P: Permutation matrix / L,U: triangle matrix)
vector([3, 1, 2])              # define vector

var('x, y')                    # define variables
plot3d(y^2+1-x^3-x, (x, -pi, pi), (y, -pi, pi))
                               # 3D Plot
implicit_plot3d(n.inner_product(p_0-p)==0, (x,
-10, 10 , (y, -10, 10), (z, -10, 10))
                               # 3D Hyperplane Plot
var('t')  # define variable (parametric equations)
x=2+2*t
y=-3*t-2
parametric_plot((x,y), (t, -10, 10), rgbcolor='red')
                               # line Plot
```

I. (3pt x 6= 18pt) True(T) or False(F). Let $A \in M_{n\times n}$ and $\mathbf{u}, \mathbf{v} \in R^n$.

1. (   ) Every square matrix can be expressed as products of elementary matrices.

2. (   ) Let $n\times n$ matrix A has all integer components. If the determinant of A is 1, then the components of A^{-1} are all integers.

3. (   ) Let $A, B \in M_n$, then $\det(AB) = \det(BA)$

4. (   ) $\mathrm{proj}_{\mathbf{u}}\mathbf{v} = \dfrac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$

5. (   ) One can compute the solution of a system of linear equations with n unknowns and n equations by Cramer's rule.

6. (   ) 2×2 real matrix A satisfies $\lambda^2 + \mathrm{tr}(A)\lambda + \det(A) = 0$.

II. (3pt x 4 = 12pt) State or Define

1. Choose 4 items from the list given in the box and describe them clearly and concisely.

normal vector of a plane π, linearly independent and linearly dependent, condition for subspace, Cramer's Rule, eigenvalue, eigenvector, linear transformations, orthogonal matrix, for linear transformation's $T : R^n \to R^m$ range, surjective or onto, injective or 1-1, isomorphism

[Subspace] A nonempty subset W of R^n satisfying the following two properties,

$$\begin{cases} \mathrm{w}_1 + \mathrm{w}_2 \in W & (1) \\ k\mathrm{w} \in W & (2) \end{cases}$$

is called a subspace of R^n. (where, $\mathrm{w}_1, \mathrm{w}_2, \mathrm{w} \in W$, $k \in R$)

...

[Standard Matrix] For a linear transformation, $T : R^n \to R^m$ the range is defined as

$$\mathrm{Im}\, T = \{ T(\mathrm{v}) \in R^m \ : \mathrm{v} \in R^n \} \subset R^m.$$

If $T : R^n \to R^m$ is a linear transformation and $A = [\,T\,]$ is the standard matrix of T, then for $\mathrm{x} \in R^n$, $T(\mathrm{x}) = A\mathrm{x}$, $\forall\ \mathrm{x} \in R^n$ where $A = [\, T(\mathrm{e}_1) : T(\mathrm{e}_2) : \cdots : T(\mathrm{e}_n) \,]$. ...

III. (3pt x 10 = 30pt) Find or Explain:

...

2. Find the equation of a plane passing through a point $P(10, -15, 4)$ and generated by two vectors $\mathrm{a} = (4, 8, 7)$ and $\mathrm{b} = (4, 5, -6)$in a vector equation form.

Ans
$$\begin{aligned} \mathrm{x} &= \mathrm{p} + t_1 \mathrm{a} + t_2 \mathrm{b} \quad (t_1,\ t_2 \in R) \\ &= (10, -15, 4) + t_1 (4, 8, 7) + t_2 (4, 5, -6). \end{aligned}$$ ■

...

5. Suppose you got a job in a research lab and your boss asked you to find the eigenvalues, the eigenvectors, and the characteristic polynomial of a matrix $A = \begin{bmatrix} 4 & 1 & 0 & 2 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 4 & 0 & 3 \end{bmatrix}$. Explain how to find them with a step by step description. You may use Sage.

Sol

1) Step 1: Open the webpage http://math1.skku.ac.kr .
2) Step 2: Log in to the webpage with ID= skku, PW = *** .
3) Step 3: Press the button of "New Worksheet"
4) Step 4: In the first cell, define matrix A in CDF format.
   A=matrix(CDF,4,4,[4,1,0,2,0,-1,2,0,0,0,1,0,0,4,0,3])
5) Step 5: In the second cell, enter the command to find eigenvalues
   A.eigenvalues() and execute.
6) Step 6: In the third cell, enter the command to find eigenvectors
   A.eigenvectors_right() and execute.
7) Step 7: In the fourth cell, enter the command to find characteristic polynomial
   A.charpoly() and execute.

[4.0, 3.0, -1.0, 1.0]

[(4.0, [(**1.0, 0, 0, 0)**], 1), (3.0, [(**0.894427191, 0, 0, -0.4472135955)**], 1), (-1.0, [(**0.140028008403, 0.700140042014, 0, -0.700140042014)**], 1),
(1.0, [(**-0.377964473009, -0.377964473009, -0.377964473009, 0.755928946018)**], 1)]

x^4 - 7.0*x^3 + 11.0*x^2 + 7.0*x – 12.0 **(Online Sage solution)** ■

.....

8. For a given matrix $A = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$, describe step by step process to find the inverse matrix by using the Sage.

Sol

1) Step 1: (example) Open the webpage http://math1.skku.ac.kr .
2) Step 2: Log in to the webpage with ID= skku, PW = *** .
3) Step 3: Press the button of "New Worksheet"
4) Step 4: In the first cell, define matrix A in CC format.
   A=matrix(CC, 3, 3, [1,0,1,-0,3,0,1,0,2])
5) Step 5: In the second cell, enter the command **A.inverse()** to find the inverse.

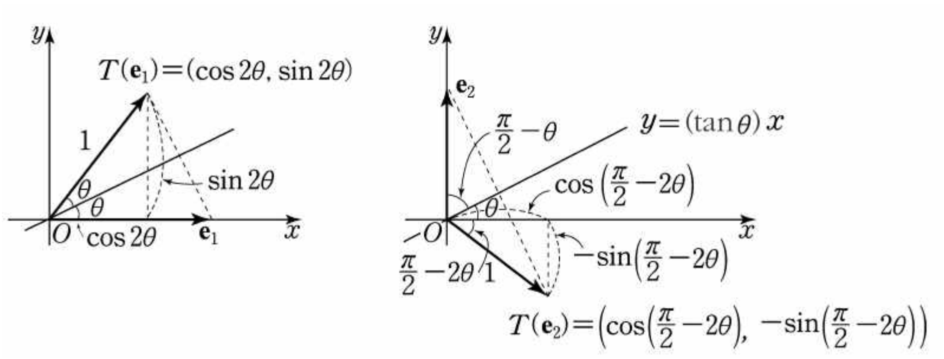
[2.00 0 -1.0]
[0 0.33 0]
[-1.00 0 1.00] ■

...

9. ...

IV. (5pt x 5 = 25pt) Find or Explain:

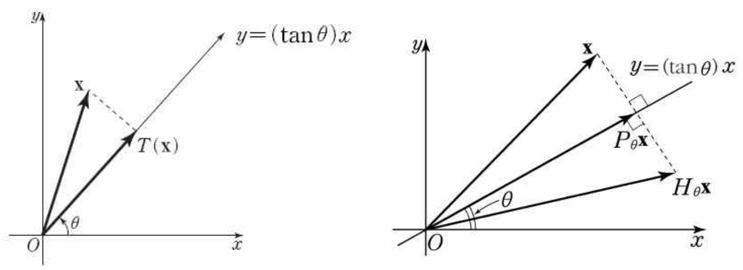
1. Let a linear transformation $T : R^2 \to R^2$ transforms any vector $\mathrm{x}=(x,y)\in R^2$ to a symmetric point to the line which passing through the origin with slope θ. Find the transformation matrix $H_\theta=[T(\mathrm{e}_1) : T(\mathrm{e}_2)]$ with the aid of following pictures.



Picture: The image of the standard basis by a symmetric transformation to the line with slope θ.

(Sol) $H_\theta=[T(\mathrm{e}_1) : T(\mathrm{e}_2)]=\begin{bmatrix}\cos 2\theta & \cos\left(\frac{\pi}{2}-2\theta\right)\\ \sin 2\theta & -\sin\left(\frac{\pi}{2}-2\theta\right)\end{bmatrix}=\begin{bmatrix}\cos 2\theta & \sin 2\theta\\ \sin 2\theta & -\cos 2\theta\end{bmatrix}$. ■

2. Linear transformation (Linear operator): Let's define $T : R^2 \to R^2$ as a projective transformation, which transforms any vector x in R^2 to projection on a line which passes through the origin and has an angle θ with x-axis. For the given transformation T, let's define P_θ as a corresponding standard matrix. As shown by the right hand side picture, $P_\theta\mathrm{x}-\mathrm{x}=\frac{1}{2}(H_\theta\mathrm{x}-\mathrm{x})$ <same direction with half length>. Now by using the matrix representation of symmetric transformation $H_\theta=\begin{bmatrix}\cos 2\theta & \sin 2\theta\\ \sin 2\theta & -\cos 2\theta\end{bmatrix}$, find the standard matrix for T.



Picture: Projective transformation to the line with slope θ The relationship between symmetric transformation and projective transformation to the line with slope θ

(Sol) $P_\theta\mathrm{x}-\mathrm{x}=\frac{1}{2}(H_\theta\mathrm{x}-\mathrm{x})$ => $P_\theta\mathrm{x}=\frac{1}{2}H_\theta\mathrm{x}+\frac{1}{2}\mathrm{x}=\frac{1}{2}H_\theta\mathrm{x}+\frac{1}{2}I\mathrm{x}=\frac{1}{2}(H_\theta+I)\mathrm{x}$

=> $P_\theta = \frac{1}{2}(H_\theta + I) = \begin{bmatrix} \frac{1}{2}(1+\cos 2\theta) & \frac{1}{2}\sin 2\theta \\ \frac{1}{2}\sin 2\theta & \frac{1}{2}(1-\cos 2\theta) \end{bmatrix} = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix}$ ■

3. For invertible matrices A and B, explain why $\mathrm{adj}(AB) = \mathrm{adj}\,B \cdot \mathrm{adj}\,A$.

...

4. For a degree n square matrix A, explain why its eigenspace is a subspace of R^n.

Ans For a given square matrix A, let $E(\lambda_i)$ be the eigenspace corresponding to an eigenvalue λ_i.

$E(\lambda_i) = \{\mathrm{x} \in R^n | A\mathrm{x} = \lambda_i \mathrm{x}\} \subseteq R^n$, $k \in R$, $E(\lambda_i) \neq \varnothing$

$\forall$ x, y $\in E(\lambda_i)$, x + y $\in \mathrm{R}^n$ and $k\mathrm{x} \in \mathrm{R}^n$.

1) [Show the space is closed under the addition, that is, show x + y $\in E(\lambda_i)$] **(2 pt)**

(Proof) $A\mathrm{x} = \lambda_i \mathrm{x}$, $A\mathrm{y} = \lambda_i \mathrm{y}$ $\therefore$ x + y $\in E(\lambda_i)$

$A(\mathrm{x}+\mathrm{y}) = A\mathrm{x} + A\mathrm{y} = \lambda_i \mathrm{x} + \lambda_i \mathrm{y} = \lambda_i(\mathrm{x}+\mathrm{y})$

2) [Show the space is closed under the scalar multiplication, that is, show $k\mathrm{x} \in E(\lambda_i)$] **(2 pt)**

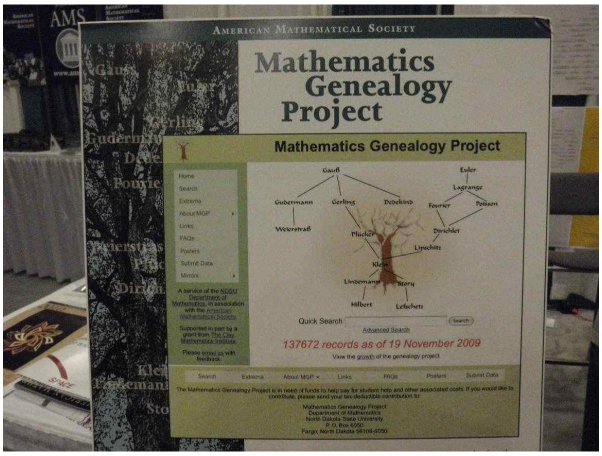
(Proof) $A(k\mathrm{x}) = kA\mathrm{x} = k\lambda_i \mathrm{x} = \lambda_i(k\mathrm{x})$

$\therefore k\mathrm{x} \in E(\lambda_i)$

$\therefore E(\lambda_i)$ is a subspace of R^n as it fulfilled the above two conditions. **(1 pt)** ■

5. For a linear transformation $T : R^n \to R^m$, explain why $\mathrm{Im}\,T$ is a subspace of R^m.

(Proof) .... ■



[Math Genealogy Tree]

http://genealogy.math.ndsu.nodak.edu

http://genealogy.math.ndsu.nodak.edu/id.ph
p?id=45061

BIGBOOK
공유와 협력의 교과서 만들기 운동본부

This book was translated by the following authors from a free **Big Book in** http://www.bigbook.or.kr/ which is an interactive smart Linear Algebra textbook for everyone.

[Authors]

http://matrix.skku.ac.kr/sglee/vita/LeeSG.htm

Sang-Gu LEE is a Professor of Mathematics at Sungkyunkwan University and Vice president of Korean Mathematical Society. He was a director of BK 21 Math Modeling Division, LOC member of ICM 2014, ILAS 2014 and ICME 12. He has held visiting appointments at the College of William and Mary, University of Northern Iowa and Utah State University. He is a member of the International Linear Algebra Society's Education Committee. His areas of interest include Matrix Theory and Mathematical Modeling, Computational Mathematics. He has authored about 20 books and published 140 scientific papers. sglee@skku.edu

Jon-Lark Kim is an Associate Professor of Mathematics at Sogang Unviersity, South Korea. He received the Ph.D. degree in mathematics from the University of Illinois at Chicago, in 2002. From 2005 to 2012 he was with the Department of Mathematics at the University of Louisville. He awarded a 2004 Kirkman Medal of the Institute of Combinatorics and its Applications. He is a member of the Editorial Board of both “Designs, Codes, and Cryptography” (2011-current) and “Journal of Algebra, Combinatorics, Discrete Structures and Applications” (2014-current). His areas of interest include Coding Theory and its interaction with Algebra, Combinatorics, Number Theory, Cryptography, and Industrial Mathematics.

In-Jae KIM is a Professor of Mathematics at Minnesota State University, Mankato. His research interests include linear algebra & matrix theory and their applications in machine learning and big data. He was an assistant managing editor of Electronic Journal of Linear Algebra from Fall 2010 to Spring 2013, and has been a reviewer for mathematics journals such as Linear Algebra and its Applications, Linear and Multilinear Algebra, and Electronic Journal of Linear Algebra.

Namyong Lee is a Professor of Mathematics at Minnesota State University, Mankato. He received his Ph. D. degree in Mathematics from University of Minnesota. His other academic related appointments includes, research director of

Minnesota Modeling and Simulation Center, secretary of Mathematical Association of America, NCS, and associate editor of Frontiers in Systems Biology. His areas of interest include, Mathematical Biology, Industrial Mathematics, and Partial Differential Equations.

Ajit Kumar : Dr. Ajit Kumar is a faculty at the Department of Mathematics, Institute of Chemical Technology, Mumbai, India since 2004. He received his Ph.D. degree from the University of Mumbai in 2002. His main interests are differential geometry, optimization techniques and the use of technology in teaching mathematics. He has initiated a lot of mathematicians into the use of open source mathematics software. He has authored two books and several book chapters. For the last several years, Dr. Ajit Kumar have been associated with the Mathematics Training and Talent Search Programme, India, aimed at undergraduates.

Victoria Lang is a graduate student of Mathematics at Sungkyunkwan University in South Korea. Originally from Maine, USA, she received her Bachelor's of Science in Mathematics and a minor in Korean Language & Culture from the University of Pittsburgh. Her most recent research interests are mathematics with 3-D printers, integrating 3-D printers into the curriculum of undergraduate-level courses, and mathematical modelling. She currently is the webmaster of the American International Student Association in Korea.

Phong Vu is a Professor of Mathematics at Ohio University who is active in research on Operator and Matrix Analysis. He got his Dr.Sc., degree from Kharkov State University, Russia. Recently, Sang-Gu and Phong wrote several interesting papers together.

Jae Hwa Lee received his Ph.D. from Chinese Academy of Sciences in 2009, And he was a Postdoctoral Researcher at BK21 Math. Modeling HRD. Div., Sungkyunkwan University from 2010-2012. Now he is a Lecturer in the Department of Mathematics at Hallym University. His research interests include Computational Mathematics, Mathematics Education and History of Korean Mathematics, etc.

< The End >

BIGBOOK
공유와 협력의 교과서 만들기 운동본부

http://sage.skku.edu http://www.sagemath.org

http://matrix.skku.ac.kr/LA-Lab/

"Headquarter for sharing free textbooks" was made to share knowledges and university textbook for educational equality and a better future of our society.
http://bigbook.or.kr/

9 788924 015249
93410
ISBN 978-89-24-01524-9