MA1101R Tutorial

Xiang Sun[1,2]

Department of Mathematics

April 15, 2011

[1] Email: xiangsun@nus.edu.sg

[2] Corrections are always welcome.

Schedule of Tutorial 1

- Review concepts:
  - Linear equation, Linear system;
  - Elementary Row Operations (ERO), Gaussian Elimination (GE) and Gauss-Jordan Elimination (GJE);
  - Row-Echelon Form (REF), Reduced Row-Echelon Form (RREF).
- Tutorial: 1.8, 1.13, 1.18(b), 1.21, 1.22, 1.23
- Additional material:
  - Structure theorem for the solutions of the linear systems
  - Question 1 in Final of 2001–2002(II)
  - Question 1 in Final of 2003–2004(II)

History of Linear Systems

- About 4000 years ago the Babylonians knew how to solve a system of two linear equations in two unknowns (a 2×2 system);
- In the famous *Nine Chapters on the Mathematical Art* (九章算术) (c. 200 BC), the Chinese solved 3×3 systems by working solely with their (numerical) coefficients;
- The modern study of systems of linear equations can be said to have originated with Leibniz[3], who in 1693 invented the notion of a determinant (Def 2.5.2) for this purpose;
- In *Introduction to the Analysis of Algebraic Curves* of 1750, Cramer[4] published the rule (Thm 2.5.32) named after him for the solution of an $n \times n$ system;
- Euler[5] was perhaps the first to observe that a system of n equations in n unknowns does not necessarily have a unique solution;
- About 1800, Gauss[6] introduced a systematic procedure, now called Gaussian Elimination, for the solution of systems of linear equations, though he did not use the matrix notation.

---

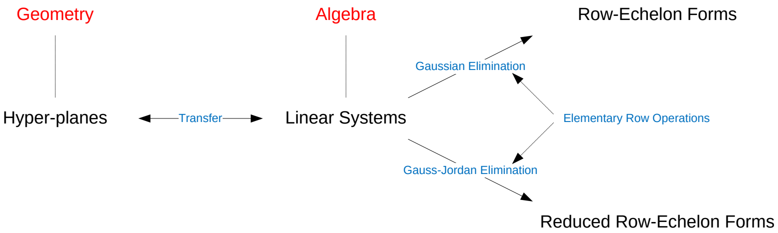
[3] Gottfried Leibniz (July 1, 1646–November 14, 1716), a German mathematician and philosopher.

[4] Gabriel Cramer (July 31, 1704–January 4, 1752), a Swiss mathematician.

[5] Leonhard Euler (April 15, 1707–September 18, 1783), a pioneering Swiss mathematician and physicist.

[6] Carl Friedrich Gauss (April 30, 1777–February 23, 1855), a German mathematician and scientist.

Structure of Chapter 1



Key:

1. Studying some geometric problems = Studying the relative linear system. For example, solving a linear system = Finding the intersection of the graphs of the equations in this linear system;
2. If the augmented matrix of a linear system is in REF or RREF, we can get the solutions easily.

Transfer between Linear Systems and Hyper-planes

A linear equation is an (algebraic) equation in which each term is either a constant or the product of a constant and (the first power of) a single variable.

Dimen	Geometric view	Algebraic representation
2	points on a line	solutions of $ax + by = c$
3	points on a plane	solutions of $ax + by + cz = d$
$n(>3)$	points on a hyper-plane	solutions of $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$
2	intersection of 2 lines	solutions of the system $\begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{cases}$
3	intersection of 2 planes	solutions of the system $\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_1 \end{cases}$
$n(>3)$	intersection of 2 hyper-planes	solutions of the system $\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \end{cases}$
2	intersection of m(>1) lines	solutions of the system $\begin{cases} a_1x + b_1y = c_1 \\ \cdots\cdots\cdots\cdots \\ a_mx + b_my = c_m \end{cases}$
3	intersection of m(>1) planes	solutions of the system $\begin{cases} a_1x + b_1y + c_1z = d_1 \\ \cdots\cdots\cdots\cdots \\ a_mx + b_my + c_mz = d_m \end{cases}$
$n(>3)$	intersection of m(>1) hyper-planes	solutions of the system $\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$

Elementary Row Operations

- $$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_1 \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \Rightarrow \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$
- Elementary row operations (ERO):
  - Multiply a row by a nonzero constant;
  - Interchange two rows;
  - Add a multiple of one row to another row.
- Two augmented matrices are said to be row equivalent if one can be obtained from the other by a series of elementary row operations.
- Theorem 1.2.7: If augmented matrices of two systems of linear equations are row equivalent, then the two systems have the same set of solutions.
- Why perform elementary row operations: the augmented matrices will be reduced to be in REF or RREF via ERO, which is easier to solve.

Gaussian Elimination and Gauss-Jordan Elimination

- There are standard procedures to get REF and RREF, which are Gaussian elimination and Gauss-Jordan elimination, respectively.
- Gaussian Elimination:
  (1) Locate the leftmost column that does not consist entirely of zeros;
  (2) Interchange the top row with another row, if necessary, to bring a nonzero entry to the top of the column found in Step 1.
  (3) For each row below the top row, add a suitable multiple of the top row to it so that the entry below the leading entry of the top row becomes zero.
  (4) Now cover the top row in the matrix and begin again with Step 1 applied to the submatrix that remains. Continuous in this way until the entire matrix is in row-echelon form.
- Gauss-Jordan Elimination: For a REF of an augmented matrix, use Gauss-Jordan elimination to reduce it to be in RREF:
  (5) Multiple a suitable constant to each row so that all the leading entries become 1.
  (6) Beginning with the last nonzero row and working upward, add suitable multiples of each row to the rows above to introduce zeros above the leading entries.

Row-Echelon Form and Reduced Row-Echelon Form

- An augmented matrix is said to be in row-echelon form (REF) if it has properties 1 and 2:
  (1) If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.
  (2) In any two successive nonzero rows, the first nonzero number in the lower row occurs farther to the right than the first nonzero number in the higher row.
- Definitions:
  - In a REF, every first nonzero number in a row is called the leading entry of the row.
  - In a REF, the leading entries of nonzero rows are also called pivot points.
  - A column of a REF is called a pivot column if it contains a pivot point; otherwise, it is called a non-pivot column.
  - Theorem: In a REF, (# nonzero rows) = (# leading entries) = (# pivot columns) = (# pivot points).
- An augmented matrix is said to be in reduced row-echelon form (RREF) if it is has properties 1, 2, 3 and 4:
  (3) The leading entry of every nonzero row is 1.
  (4) In each pivot column, except the pivot point, all other entries are zeros.

Structure Theorem for Solutions—Remark 1.4.7

A linear system has no solution if and only if the last column of its REF of the augmented matrix is a pivot column, i.e. there is a row with nonzero last entry but zero elsewhere.

$$\left(\begin{array}{ccccc|c}
 & \otimes & * & & & * \\
 & & \otimes & * & & * \\
 & 0 & & \ddots & & * \\
 & & & & \otimes \quad * & * \\
\hline
0 & \cdots & & \cdots & 0 & \otimes \\
\hline
 & & & & &
\end{array}\right) \leftarrow \text{nonzero}$$

where each $\otimes$ represents a pivot point (the leading entry of a nonzero row).

Structure Theorem for Solutions—Remark 1.4.7 (Cont.)

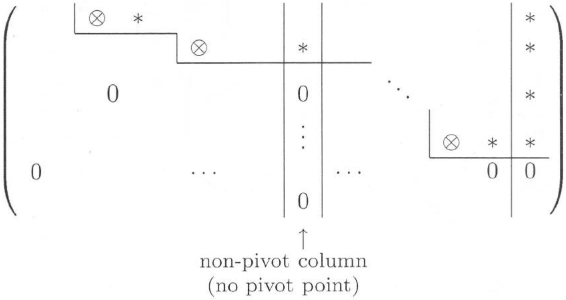
- A linear system has exactly one solution if and only if except the last column, every column of a REF of the augmented matrix is a pivot column.
- That is, A linear system has exactly one solution if and only if it is consistent and (# variables) $=$ (# nonzero rows).
- In this case, its general solution has # variables $-$ # nonzero rows $= 0$ arbitrary parameter.

$$\left(\begin{array}{cccc|c} \otimes & & & & * \\ & \otimes & & & * \\ & 0 & \ddots & & * \\ & & & \otimes & * \\ 0 & & \cdots & 0 & 0 \end{array}\right)$$

where each $\otimes$ represents a pivot point (the leading entry of a nonzero row).

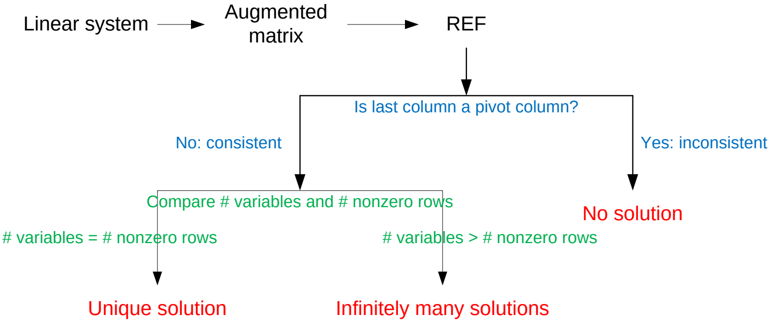
Structure Theorem for Solutions—Remark 1.4.7 (Cont.)

- A linear system has infinitely many solutions if and only if apart from the last column, a REF of the augmented matrix has at least one more non-pivot column.
- That is, A linear system has exactly one solution if and only if it is consistent and (# variables) $>$ (# nonzero rows).
- In this case, its general solution has $\#$ variables $-$ $\#$ nonzero rows $\neq 0$ arbitrary parameter(s).



where each $\otimes$ represents a pivot point (the leading entry of a nonzero row).

Structure Theorem for Solutions—Summary



Exercise (1.8)

Each equation in the following linear system represents a line in the xy-plane

$$\begin{cases} a_1 x + b_1 y = c_1 \\ a_2 x + b_2 y = c_2 \\ a_3 x + b_3 y = c_3 \end{cases}$$

where $a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3$ are constants. Discuss the relative positions of the three lines when the system

(a) *has no solution;*

(b) *has exactly one solution;*

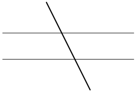
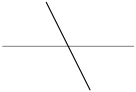
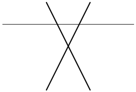
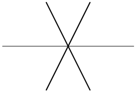
(c) *has infinitely many solutions.*

Recall

There is a one-to-one correspondence between the solution set of the linear system and the intersection of all the three lines.

Method

Let's consider all possible cases of the relative positions of 3 lines in xy-plane:

3 parallel lines:	 i. All are distinct	 ii. 2 parallel lines are same	iii. All are same
2 parallel lines:	 iv. 2 parallel lines are distinct	 v. 2 parallel lines are same	
No parallel line:	 vi. No intersection point	 vii. 1 intersection point	

Solution.

Based on the graphs above, we have the following results:

(a) In case i, ii, iv and vi, the system has no solution;

(b) In case v and vii, the system has exactly one solution;

(c) In case iii, the system has infinitely many solutions.

To summarize,

(a) When the system has no solution, either (i) the three lines are parallel but not all three are the same or (ii) two of the lines intersect at a point but this point does not lie on the third line.

(b) When the system has exactly one solution, all three lines are distinct and intersect at a single point, or two of the lines are identical and they intersect the third line at a point.

(c) When the system has infinitely many solutions, all three lines are identical.

$\square$

Exercise (1.13)

Solve the following system of non-linear equations:

$$\begin{cases} x^2 - y^2 + 2z^2 = 6 \\ 2x^2 + 2y^2 - 5z^2 = 3 \\ 2x^2 + 5y^2 + z^2 = 9 \end{cases}$$

Method

The given system of equations is not linear. Consider replacing the variables x^2, y^2, z^2 by another set of variables so that we obtain a linear system.

Solution.

1. Let $u = x^2$, $v = y^2$, and $w = z^2$, then the system becomes

$$\begin{cases} u - v + 2w = 6 \\ 2u + 2v - 5w = 3 \\ 2u + 5v + w = 9 \end{cases} \Rightarrow \left(\begin{array}{ccc|c} 1 & -1 & 2 & 6 \\ 2 & 2 & -5 & 3 \\ 2 & 5 & 1 & 9 \end{array}\right)$$

2. Apply Gauss-Jordan elimination to obtain the RREF:

$$\to \left(\begin{array}{ccc|c} 1 & -1 & 2 & 6 \\ 0 & 4 & -9 & -9 \\ 0 & 7 & -3 & -3 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -1 & 2 & 6 \\ 0 & -17 & 0 & 0 \\ 0 & 7 & -3 & -3 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -1 & 2 & 6 \\ 0 & 1 & 0 & 0 \\ 0 & 7 & -3 & -3 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|c} 1 & 0 & 2 & 6 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -3 & -3 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 2 & 6 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

Then $x^2 = u = 4$, $y^2 = v = 0$ and $z^2 = w = 1$.

3. Thus the solutions are $(x, y, z) = (2, 0, 1), (2, 0, -1), (-2, 0, 1), (-2, 0, -1)$.

$\square$

Exercise (1.18(b))

For

$$\begin{cases} x + y + z = 1 \\ 2x + ay + 2z = 2 \\ 4x + 4y + a^2 z = 2a \end{cases}$$

determine the values of a such that the system has

(i) *no solution;*

(ii) *exactly one solution;*

(iii) *infinitely many solutions.*

Solution.

The augmented matrix and its REF are:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & a & 2 & 2 \\ 4 & 4 & a^2 & 2a \end{array}\right) \xrightarrow[R_3 - 4R_1]{R_2 - 2R_1} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & a-2 & 0 & 0 \\ 0 & 0 & a^2-4 & 2(a-2) \end{array}\right)$$

Therefore, the system has no solution if $a = -2$. It has only one solution if $a \neq 2, -2$. It has infinitely many solutions if $a = 2$. $\square$

Exercise (1.21)

Consider the homogeneous linear system

$$\begin{cases} a_1x + b_1y + c_1z = 0 \\ a_2x + b_2y + c_2z = 0 \\ a_3x + b_3y + c_3z = 0 \end{cases}$$

where $a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3$ *are constants. Determine all possible reduced row-echelon forms of the augmented matrix of the system and describe the geometrical meaning of the solutions obtained from various reduced row-echelon forms.*

Remark

For a linear system, the coefficients may be zeros.

Solution.

- A reduced-row echelon form with three nonzero rows (leading entries):

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

  Since homogeneous system is always consistent, and $\#$ variables $= 3 = \#$ nonzero rows, the solution is unique, i.e., the origin in $\mathbb{R}^3$.

- Reduced-row echelon forms with two nonzero rows (leading entries):

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

  Since homogeneous system is always consistent, and $\#$ variables $= 3 > 2 = \#$ nonzero rows, the general solutions obtained here has one parameter. Thus, the solutions represent lines in $\mathbb{R}^3$ that passes through the origin.

$\square$

Solution (Cont.)

- Reduced-row echelon forms with one nonzero row (leading entry):

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

  Since homogeneous system is always consistent, and $\#$ variables $= 3 > 1 = \#$ nonzero rows, the general solutions obtained here has two parameter. Thus, the solutions obtained here represent planes in $\mathbb{R}^3$ that passes through the origin.

- A reduced-row echelon form with zero nonzero row (leading entry):

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

  The solutions obtained here represent the whole space of $\mathbb{R}^3$.

□

Exercise (1.22)

Let $\left(\begin{array}{ccc|c} a & 0 & 0 & d \\ 0 & b & 0 & e \\ 0 & 0 & c & f \end{array}\right)$ *be the reduced row-echelon form of the augmented matrix of a linear system, where* a, b, c, d, e, f *are real numbers. Write down the necessary conditions on* a, b, c, d, e, f *so that the solution set of the linear system is a plane in the three dimensional space that does not contain the original.*

Solution.

- For the solution set to be a plane, there must be one leading entry in the reduced-row echelon form and two arbitrary parameters. Thus, we have two zero rows, i.e. $b = c = e = f = 0$.
- Since the system is consistent (The solution set is a plane) and these is one nonzero row, a is a leading entey, which is 1.
- Since the plane does not contain the origin, $d \neq 0$.

To summarize: the necessary condition is $a = 1$, $b = c = e = f = 0$, $d \neq 0$. □

Exercise (1.23)

(a) *Does an inconsistent linear system with more unknowns than equation exist?*

(b) *Does a linear system which has exactly one solution, but more equations than unknowns, exist?*

(c) *Does a linear system which has exactly one solution, but more unknowns than equations, exist?*

(d) *Does a linear system which has infinitely many solutions, but more equations than unknowns, exist?*

Solution.

(a) Yes. For example: $\begin{cases} x+y+z=0 \\ x+y+z=1 \end{cases}$.

(b) Yes. For example: $\begin{cases} x=0 \\ 2x=0 \end{cases}$.

(c) No. A linear system with more unknowns than equations will either have no solution or infinitely many solutions.

(d) Yes. For example: $\begin{cases} x+y=0 \\ 2x+2y=0 \\ 3x+3y=0 \end{cases}$.

□

Structure theorem for the solutions of the linear systems

- The rank of a matrix (Def 4.2.3) is the dimension of its row space (or column space).
- Notation: $\mathrm{rank}(\boldsymbol{A})$.
- Theorem: $\mathrm{rank}(\boldsymbol{A})$ is equal to the # nonzero pivot columns in a REF of $\boldsymbol{A}$.
- Homogeneous:
  - $\boldsymbol{A}_{m\times n} \cdot \boldsymbol{x}_{n\times 1} = \boldsymbol{0}_{m\times 1}, \mathrm{rank}(\boldsymbol{A}) = r \leq \min\{m, n\}$;
  - The homogeneous system is always consistent;
  - If $r = n$, then there is only zero solution;
  - If $r < n$, then there are infinite solutions with $n - r$ arbitrary parameter(s).
- Inhomogeneous:
  - $\boldsymbol{A}_{m\times n} \cdot \boldsymbol{x}_{n\times 1} = \boldsymbol{b}_{m\times 1}, \mathrm{rank}(\boldsymbol{A}) = r \leq \min\{m, n\}$;
  - The inhomogeneous system is consistent iff $\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A} \mid \boldsymbol{b})$.
  - If consistent, and $r = n$, then there is only one solution;
  - If consistent, and $r < n$, then there are infinite solutions with $n - r$ arbitrary parameter(s).

Exercise (Question 1 in Final of 2001–2002(II))

Find a condition on the numbers a, b and c such that the following system of equations is consistent. When that condition is satisfied, find all solutions (in terms of a and b).

$$\begin{cases} x+3y+z=a \\ -x-2y+z=b \\ 3x+7y-z=c \end{cases}$$

Solution.

Apply Gaussian elimination for the relative augmented matrix, we will obtain:

$$\left(\begin{array}{ccc|c} 1 & 3 & 1 & a \\ -1 & -2 & 1 & b \\ 3 & 7 & -1 & c \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 3 & 1 & a \\ 0 & 1 & 2 & a+b \\ 0 & -2 & -4 & c-3a \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 3 & 1 & a \\ 0 & 1 & 2 & a+b \\ 0 & 0 & 0 & -a+2b+c \end{array}\right).$$

Thus, this linear system is consistent iff $-a+2b+c=0$. □

Exercise (Question 1 in Final of 2003–2004(II))

Find a condition on the numbers a and b such that the following system of equations is not consistent.

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

Solution.
Apply Gaussian elimination for the relative augmented matrix, we will obtain:

$$\left(\begin{array}{ccc|c} 1 & 3 & 0 & a \\ 2 & 4 & 2 & 2a \\ 3 & 3b & 3 & 6 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 3 & 0 & a \\ 0 & -1 & 1 & 0 \\ 0 & 3b-9 & 3 & 6-3a \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 3 & 0 & a \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 3b-6 & 6-3a \end{array}\right).$$

Thus, this linear system is inconsistent iff $6 - 3a \neq 0$ and $3b - 6 = 0$, that is, $a \neq 2$ and $b = 2$. $\square$

Change log

- Page 12: Add an diagram for how to use Structure Theorem;
- Page 22: Revise the Solution.

Last modified: 13:24, January 28, 2011.

Schedule of Tutorial 2

- Any question about last tutorial
- Review concepts:
  - Definition of matrices;
  - Matrix operations: addition, scalar multiplication, multiplication and transpose;
  - Inverse, elementary matrices.
- Tutorial: 2.7, 2.10, 2.11, 2.15, 2.16, 2.19
- Additional material:
  - 2.9, 2.20, 2.21, 2.22;
  - Question 5 in Final of 2006–2007(I);
  - Question 2 in Final of 2001–2002(II);
  - Question 1(b) in Final of 2005–2006(I);
  - Three extra questions.

History of Matrix Theory

- Matrices were introduced implicitly as abbreviations of linear transformations by Gauss;
- Arthur Cayley[7] formally introduced $m \times n$ matrices in two papers in 1850 and 1858 (the term “matrix” was coined by Sylvester[8] in 1850);
- In his 1858 paper “A memoir on the theory of matrices” Cayley proved the important Cayley–Hamilton theorem of 2×2 and 3×3 matrices, while Hamilton[9] proved the theorem independently for 4×4 matrices;
- Cayley advanced considerably the important idea of viewing matrices as constituting a symbolic algebra. But his papers of the 1850s were little noticed outside England until the 1880s;
- During 1820s–1870s, Cauchy, Jacobi, Jordan[10], Weierstrass, and others created what may be called the spectral theory of matrices; An important example is the Jordan canonical form;
- In a seminal paper in 1878 titled “On linear substitutions and bilinear forms” Frobenius[11] developed substantial elements of the theory of matrices in the language of bilinear forms.

---

[7] Arthur Cayley (August 16, 1821–January 26, 1895), a British mathematician.

[8] James Joseph Sylvester (September 3, 1814–March 15, 1897), an English mathematician.

[9] William Rowan Hamilton (August 4, 1805–September 2, 1865), an Irish physicist, astronomer, and mathematician.

[10] Camille Jordan (January 5, 1838–January 22, 1922), a French mathematician.

[11] Ferdinand Georg Frobenius (October 26, 1849–August 3, 1917), a German mathematician.

Definition and Notation

- A matrix is a rectangular array of numbers (the numbers here can be in $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$, etc.). The size of a matrix is given by $m \times n$ where m is # rows and n is # columns. The (i, j)-entry of a matrix is the number which is in the ith row and jth column of the matrix.
- In general, an $m \times n$ matrix can be written as

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

  or simply $\boldsymbol{A} = (a_{ij})_{m \times n}$ where a_{ij} is the (i, j)-entry of $\boldsymbol{A}$.
- Two matrices are said to be equal if they have the same size and their corresponding entries are equal.
- Remark: in some case, we regard (a) (a 1×1 matrix) and a (a scalar) to be same.

Special Types of Matrices (I)

- Row (resp. Column) matrix: only 1 row (resp. column);
- Square matrix: # rows = # columns;
- Diagonal matrix: square matrix, $a_{ij} = 0$ when $i \neq j$;
- Tridiagonal matrix: nonzero elements only in the main diagonal, the first diagonal below this, and the first diagonal above this;
- Identity matrix: diagonal matrix where diagonal entries are 1. Notation: $\boldsymbol{I}_n$;
- Scalar matrix: diagonal matrix where diagonal entries are c–constant number. Notation: $c\boldsymbol{I}_n$;
- Zero matrix: all entries are 0, notation: $\mathbf{0}_{m\times n}$;
- Upper (resp. Lower)-triangular matrix: square matrix, $a_{ij} = 0$ if $i > j$ (resp. $i < j$) this.
  - Multiplication of two upper (resp. lower)-triangular matrices is also an upper (resp. lower)-triangular matrix;
  - Inverse of an upper (resp. lower)-triangular invertible matrix is upper (resp. lower)-triangular.

Addition

Let $\boldsymbol{A} = (a_{ij})_{m\times n}$ and $\boldsymbol{B} = (b_{ij})_{m\times n}$. Define the matrix addition

$$\boldsymbol{A} + \boldsymbol{B} = (a_{ij} + b_{ij})_{m\times n}.$$

- Associated law: Let $\boldsymbol{A} = (a_{ij})_{m\times n}$, $\boldsymbol{B} = (b_{ij})_{m\times n}$ and $\boldsymbol{C} = (c_{ij})_{m\times n}$, then $(\boldsymbol{A}+\boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$;
- Commutative law: Let $\boldsymbol{A} = (a_{ij})_{m\times n}$ and $\boldsymbol{B} = (b_{ij})_{m\times n}$, then $\boldsymbol{A}+\boldsymbol{B} = \boldsymbol{B}+\boldsymbol{A}$;
- Identity: Let $\boldsymbol{A} = (a_{ij})_{m\times n}$, then $\boldsymbol{A} + \mathbf{0}_{m\times n} = \mathbf{0}_{m\times n} + \boldsymbol{A} = \boldsymbol{A}$;
- Inverse: For for $\boldsymbol{A} = (a_{ij})_{m\times n}$, there exists a unique matrix $\boldsymbol{B} = (b_{ij})_{m\times n}$, such that $\boldsymbol{A}+\boldsymbol{B} = \mathbf{0} = \boldsymbol{B}+\boldsymbol{A}$; We will denote $\boldsymbol{B}$ as $-\boldsymbol{A}$;
- Based on definition of $-\boldsymbol{A}$, we can define the matrix subtraction: $\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{A} + (-\boldsymbol{B})$.

Scalar Multiplication

Let $\boldsymbol{A} = (a_{ij})_{m\times n}$ and μ be a real constant. Define the scalar multiplication $\mu\boldsymbol{A} = (\mu a_{ij})_{m\times n}$, where μ is usually called a scalar.

- Let $\boldsymbol{A} = (a_{ij})_{m\times n}$ and μ, λ be real constants, then $(\mu\lambda)\boldsymbol{A} = \mu(\lambda\boldsymbol{A})$;
- $1\boldsymbol{A} = \boldsymbol{A}$;
- 1st distributive law: Let $\boldsymbol{A} = (a_{ij})_{m\times n}$ and μ, λ be real constants, then
$$(\mu + \lambda)\boldsymbol{A} = \mu\boldsymbol{A} + \lambda\boldsymbol{A};$$
- 2nd distributive law: Let $\boldsymbol{A} = (a_{ij})_{m\times n}$, $\boldsymbol{B} = (b_{ij})_{m\times n}$ and μ a be real constant, then
$$\mu(\boldsymbol{A} + \boldsymbol{B}) = \mu\boldsymbol{A} + \mu\boldsymbol{B};$$
- Let $\boldsymbol{A} = (a_{ij})_{m\times n}$ and μ be a real constant, if $\mu\boldsymbol{A} = \boldsymbol{0}$, then $\boldsymbol{A} = \boldsymbol{0}$ or $\mu = 0$.

Multiplication

Let $\boldsymbol{A} = (a_{ij})_{m\times p}$ and $\boldsymbol{B} = (b_{ij})_{p\times n}$. Define the matrix multiplication $\boldsymbol{AB} = (c_{ij})_{m\times n}$, where

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj},$$

for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

- Associated law: Let $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ be $m\times p$, $p\times q$ and $q\times n$ matrices respectively, then $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$; Moreover, we can define $\boldsymbol{A}^n = \begin{cases} \boldsymbol{I} & \text{if } n = 0 \\ \overbrace{\boldsymbol{AA}\cdots\boldsymbol{A}}^{n \text{ times}} & \text{if } n \in \mathbb{N} \end{cases}$.
- Commutative law: not always hold.
- Identity: Let $\boldsymbol{A} = (a_{ij})_{m\times n}$, then $\boldsymbol{AI}_n = \boldsymbol{I}_m\boldsymbol{A} = \boldsymbol{A}$;
- Inverse: not invertible for all matrices;
- 1st-Distributive law: Let $\boldsymbol{A}$, $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ be $m\times p$, $p\times n$ and $p\times n$ matrices respectively, then $\boldsymbol{A}(\boldsymbol{B}_1 + \boldsymbol{B}_2) = \boldsymbol{AB}_1 + \boldsymbol{AB}_2$;
- 2nd-Distributive law: Let $\boldsymbol{A}$, $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ be $p\times n$, $m\times p$ and $m\times p$ matrices respectively, then $(\boldsymbol{C}_1 + \boldsymbol{C}_2)\boldsymbol{A} = \boldsymbol{C}_1\boldsymbol{A} + \boldsymbol{C}_2\boldsymbol{A}$;
- Let $\boldsymbol{A} = (a_{ij})_{m\times p}$, $\boldsymbol{B} = (b_{ij})_{p\times n}$ and μ be a real constant, then $(\mu\boldsymbol{A})\boldsymbol{B} = \boldsymbol{A}(\mu\boldsymbol{B}) = \mu(\boldsymbol{AB})$;

Transpose

The transpose of a matrix $\boldsymbol{A}$, denoted by $\boldsymbol{A}^T$, is the matrix obtained from $\boldsymbol{A}$ by changing columns to rows, and rows to columns.

- Let $\boldsymbol{A}$ be a matrix, then $(\boldsymbol{A}^T)^T = \boldsymbol{A}$;
- Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two $m \times n$ matrices, then $(\boldsymbol{A}+\boldsymbol{B})^T = \boldsymbol{A}^T + \boldsymbol{B}^T$;
- Let $\boldsymbol{A}$ be a matrix, and μ be a scalar, then $(\mu\boldsymbol{A})^T = \mu\boldsymbol{A}^T$;
- Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be $m \times n$ and $n \times p$ matrices, respectively, then $(\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{B}^T\boldsymbol{A}^T$;

Inverse

Let $\boldsymbol{A}$ be a square matrix of order n. Then $\boldsymbol{A}$ is said to be invertible if there exists a square matrix $\boldsymbol{B}$ of order n such that $\boldsymbol{AB} = \boldsymbol{I}_n$ and $\boldsymbol{BA} = \boldsymbol{I}_n$. Such a matrix $\boldsymbol{B}$ is called an inverse of $\boldsymbol{A}$, denoted as $\boldsymbol{A}^{-1}$. A square matrix is called singular if it has no inverse.

- Uniqueness: If $\boldsymbol{B}$ and $\boldsymbol{C}$ are inverses of a square matrix $\boldsymbol{A}$, then $\boldsymbol{B} = \boldsymbol{C}$;
- Thm 2.4.5: $\boldsymbol{A}$ is invertible iff $\boldsymbol{Ax} = \boldsymbol{0}$ has trivial solution iff RREF of $\boldsymbol{A}$ is identity matrix iff $\boldsymbol{A}$ can be expressed as a product of elementary matrices;
- Let $\boldsymbol{A}$ be an invertible matrix and μ a nonzero scalar, then $(\mu\boldsymbol{A})^{-1} = \frac{1}{\mu}\boldsymbol{A}^{-1}$;
- Let $\boldsymbol{A}$ be an invertible matrix, then $(\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T$;
- Let $\boldsymbol{A}$ be an invertible matrix, then $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$;
- Let $\boldsymbol{A}$, $\boldsymbol{B}$ be two invertible matrices of the same size, then $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$;
- Let $\boldsymbol{A}$ be an invertible matrix and n be a positive integer, then we can define $\boldsymbol{A}^{-n} = (\boldsymbol{A}^{-1})^n = \underbrace{\boldsymbol{A}^{-1}\boldsymbol{A}^{-1}\cdots\boldsymbol{A}^{-1}}_{n \text{ times}}$.

Special Types of Matrices (II)

- Symmetric matrix: $\boldsymbol{A} = \boldsymbol{A}^T$;
- Skew-symmetric matrix: $\boldsymbol{A} = -\boldsymbol{A}^T$;
- Hermite matrix: $\boldsymbol{A} = \bar{\boldsymbol{A}}^T$;
  - Let $\boldsymbol{A}$ be a square matrix, then $\boldsymbol{A} + \boldsymbol{A}^T$ is a symmetric matrix, and $\boldsymbol{A} - \boldsymbol{A}^T$ is a skew-symmetric matrix;
  - Each square matrix $\boldsymbol{A}$ can be uniquely decomposed as an addition of a symmetric matrix $\boldsymbol{S}$ and a skew-symmetric matrix $\boldsymbol{K}$.
- Nilpotent matrix: $\boldsymbol{A}^k = 0$ for some positive integer k;
  - Let $\boldsymbol{A}$ be a matrix with $\boldsymbol{A}^k = 0$, then $(\boldsymbol{I} - \boldsymbol{A})^{-1} = \boldsymbol{I} + \boldsymbol{A} + \cdots + \boldsymbol{A}^{k-1}$.
- Idempotent matrix: $\boldsymbol{A}^2 = \boldsymbol{A}$;
  - Let $\boldsymbol{A}$ be an idempotent matrix, then $(\boldsymbol{I} + \boldsymbol{A})^{-1} = \frac{1}{2}(2\boldsymbol{I} - \boldsymbol{A})$;
  - $\boldsymbol{A}$ is an idempotent matrix iff $(\boldsymbol{I} - 2\boldsymbol{A})^{-1} = \boldsymbol{I} - 2\boldsymbol{A}$.

Elementary Matrices: Multiply a row by a constant

$$m \text{ rows} \left\{ \begin{array}{r} \\ \\ i\text{th row} \rightarrow \\ \\ \\ \\ \end{array} \right. \overbrace{\begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & & \cdots & \cdots & 0 \\ \vdots & & c & & & \vdots \\ \vdots & & & 1 & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & 1 \end{pmatrix}}^{m \text{ columns}} \qquad \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & & \cdots & \cdots & 0 \\ \vdots & & \frac{1}{c} & & & \vdots \\ \vdots & & & 1 & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & 1 \end{pmatrix}$$

Multiply ith row by a constant c

Its inverse
Multiply ith row by a constant $1/c$

Elementary Matrices: Interchange two rows

$$
\begin{array}{cc}
 & \begin{matrix} & & i\text{th col} & & j\text{th col} & & \end{matrix} \\
\begin{matrix} \\ \\ i\text{th row} \\ \\ j\text{th row} \\ \\ \\ \end{matrix} & \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}
\end{array}
\qquad
\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}
$$

Interchange ith row and jth row

Its inverse
Interchange ith row and jth row

Elementary Matrices: Add a multiple of a row by a constant

$$\begin{array}{rc}
 & \begin{matrix} & & i\text{th col} & & j\text{th col} & & \end{matrix} \\
\begin{matrix} \\ \\ i\text{th row} \\ \\ j\text{th row} \\ \\ \\ \end{matrix} & \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & c & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}
\end{array}
\qquad
\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & -c & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

Add c multiple of jth row on ith row

Its inverse
Add -c multiple of jth row on ith row

Elementary Matrices: Add a multiple of a row by a constant (Cont.)

$$\begin{array}{cc}
 & \begin{array}{ccccccc} & & i\text{th col} & & j\text{th col} & & \end{array} \\
\begin{array}{c} \\ \\ i\text{th row} \\ \\ j\text{th row} \\ \\ \\ \end{array} &
\begin{pmatrix}
1 & & & & & & \\
 & \ddots & & & & & \\
 & & 1 & & & & \\
 & & & \ddots & & & \\
 & & c & & 1 & & \\
 & & & & & \ddots & \\
 & & & & & & 1
\end{pmatrix}
\end{array}
\qquad
\begin{pmatrix}
1 & & & & & & \\
 & \ddots & & & & & \\
 & & 1 & & & & \\
 & & & \ddots & & & \\
 & & -c & & 1 & & \\
 & & & & & \ddots & \\
 & & & & & & 1
\end{pmatrix}$$

Add *c* multiple of *i*th row on *j*th row

Its inverse
Add -*c* multiple of *i*th row on *j*th row

Exercise (2.7)

Give an example of a 2×3 *matrix* $\boldsymbol{A}$ *such that the solution set of the linear system* $\boldsymbol{Ax} = \mathbf{0}$ *is the plane* $\{(x, y, z) \mid 2x + 3y - z = 0\}$*.*

Solution.

1. The solution set of a homogeneous linear system (represented by $\boldsymbol{Ax} = \mathbf{0}$) in the xyz-space represents the set of points that satisfies every linear equation in the linear system.
2. In this case, the solution set is $\{(x, y, z) \mid 2x + 3y - z = 0\}$, so there is only one equation to satisfy.
3. While $\boldsymbol{A}$ is a 2×3 matrix (means there are 2 equations to satisfy), so we need to "construct" the other equation which has no effect on the solution set.
4. So the matrix $\boldsymbol{A}$ can be $\begin{pmatrix} 2 & 3 & -1 \\ 0 & 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 2 & 3 & -1 \\ 2 & 3 & -1 \end{pmatrix}$, etc.

□

Exercise (2.10)

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be $m \times n$ and $n \times p$ matrices respectively.

(a) *Suppose the homogeneous linear system $\boldsymbol{Bx} = \boldsymbol{0}$ has infinitely many solutions. How many solutions does the system $\boldsymbol{ABx} = \boldsymbol{0}$ have?*

(b) *Suppose $\boldsymbol{Bx} = \boldsymbol{0}$ has only the trivial solution. Can we tell how many solutions are there for $\boldsymbol{ABx} = \boldsymbol{0}$?*

Solution.

(a) Let $\boldsymbol{x} = \boldsymbol{u}$ be any solution to the system $\boldsymbol{Bx} = \boldsymbol{0}$. Then $\boldsymbol{ABu} = \boldsymbol{A0} = \boldsymbol{0}$. The system $\boldsymbol{ABx} = \boldsymbol{0}$ has at least as many solutions as the system $\boldsymbol{Bx} = \boldsymbol{0}$. Thus it has infinitely many solutions.

(b) No. For example, let $\boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and consider the following two cases:

(i) If $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, then both $\boldsymbol{Bx} = \boldsymbol{0}$ and $\boldsymbol{ABx} = \boldsymbol{0}$ have only trivial solution;

(ii) If $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, then both $\boldsymbol{Bx} = \boldsymbol{0}$ has only trivial solution, but $\boldsymbol{ABx} = \boldsymbol{0}$ has infinitely many solutions.

$\square$

Exercise (2.11)

Let $\boldsymbol{A} = (a_{ij})_{n\times n}$ *be a square matrix. The* trace *of* $\boldsymbol{A}$*, denoted by* $\operatorname{tr}(\boldsymbol{A})$*, is defined to be the sum of the entries on the diagonal of* $\boldsymbol{A}$*, i.e.*

$$\operatorname{tr}(\boldsymbol{A}) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}.$$

(a) *Find the trace of each of the following square matrices.*

$$\text{(i)}\ \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad \text{(ii)}\ \begin{pmatrix} -1 & 3 & -4 \\ 2 & 4 & 1 \\ -4 & 2 & -9 \end{pmatrix}, \quad \text{(iii)}\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 1 & 3 & 5 & 0 \\ 1 & 3 & 5 & 7 \end{pmatrix}.$$

(b) *Let* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *be any square matrices of the same size, show that* $\operatorname{tr}(\boldsymbol{A}+\boldsymbol{B}) = \operatorname{tr}(\boldsymbol{A}) + \operatorname{tr}(\boldsymbol{B})$*;*

(c) *Let* $\boldsymbol{A}$ *be any square matrices and* k *a scalar, show that* $\operatorname{tr}(k\boldsymbol{A}) = k\operatorname{tr}(\boldsymbol{A})$*;*

(d) *Let* $\boldsymbol{C}$ *and* $\boldsymbol{D}$ *be* $m\times n$ *and* $n\times m$ *matrices respectively, show that* $\operatorname{tr}(\boldsymbol{CD}) = \operatorname{tr}(\boldsymbol{DC})$.

Solution and Proof.

(a) (i) The trace is $1 + 1 + 0 = 2$;
(ii) The trace is $(-1) + 4 + (-9) = -6$;
(iii) The trace is $1 + 3 + 5 + 7 = 16$.

(b) Let $\boldsymbol{A} = (a_{ij})_{n\times n}$ and $\boldsymbol{B} = (b_{ij})_{n\times n}$ be two matrices, where $n \in \mathbb{N}$. Then

$$\begin{aligned}
\operatorname{tr}(\boldsymbol{A}+\boldsymbol{B}) &= \operatorname{tr}\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \cdots & a_{nn}+b_{nn} \end{pmatrix} \\
&= (a_{11}+b_{11}) + (a_{22}+b_{22}) + \cdots + (a_{nn}+b_{nn}) \\
&= (a_{11}+a_{22}+\cdots+a_{nn}) + (b_{11}+b_{22}+\cdots+b_{nn}) \\
&= \operatorname{tr}(\boldsymbol{A}) + \operatorname{tr}(\boldsymbol{B})
\end{aligned}$$

(c) Let $\boldsymbol{A} = (a_{ij})_{n\times n}$ be a matrix, where $n \in \mathbb{N}$. Then

$$\operatorname{tr}(k\boldsymbol{A}) = \operatorname{tr}\begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{n1} & ka_{n2} & \cdots & ka_{nn} \end{pmatrix} = k(a_{11}+a_{22}+\cdots+a_{nn}) = k\operatorname{tr}(\boldsymbol{A}).$$

$\square$

Solution and Proof (Cont.)

(d)

1. Let $\boldsymbol{C} = (c_{ij})_{m\times n}$ and $\boldsymbol{D} = (d_{ij})_{n\times m}$.
2. Then the (i, i)-entry of $\boldsymbol{CD}$ is

$$c_{i1}\,d_{1i} + c_{i2}\,d_{2i} + \cdots + c_{in}\,d_{ni} = \sum_{j=1}^{n} c_{ij}\,d_{ji}.$$

Thus,

$$\operatorname{tr}(\boldsymbol{CD}) = \sum_{i=1}^{m}\Bigl(\sum_{j=1}^{n} c_{ij}\,d_{ji}\Bigr) = \sum_{j=1}^{n}\Bigl(\sum_{i=1}^{m} c_{ij}\,d_{ji}\Bigr).$$

3. But the (i, i)-entry of $\boldsymbol{DC}$ is

$$d_{i1}\,c_{1i} + d_{i2}\,c_{2i} + \cdots + d_{im}\,c_{mi}.$$

So

$$\operatorname{tr}(\boldsymbol{DC}) = \sum_{j=1}^{n}\Bigl(\sum_{i=1}^{m} c_{ij}\,d_{ji}\Bigr),$$

which is precisely the term on the right hand side above.

$\square$

Exercise (2.15)

Show that there are no matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ such that $\boldsymbol{AB} - \boldsymbol{BA} = \boldsymbol{I}$.

Proof.

1. Assume that there are matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ such that $\boldsymbol{AB} - \boldsymbol{BA} = \boldsymbol{I}$.
2. Then $\mathrm{tr}(\boldsymbol{AB} - \boldsymbol{BA}) = \mathrm{tr}(\boldsymbol{I})$.
3. By Question 2.11, we have $\mathrm{tr}(\boldsymbol{AB} - \boldsymbol{BA}) = \mathrm{tr}(\boldsymbol{AB}) - \mathrm{tr}(\boldsymbol{BA}) = 0$.
4. Since $\mathrm{tr}(\boldsymbol{I})$ is the size of $\boldsymbol{I}$ which is nonzero, there is a contradiction.
5. Thus, there are no matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ such that $\boldsymbol{AB} - \boldsymbol{BA} = \boldsymbol{I}$.

$\square$

Exercise (2.16)

Determine which of the following statements are true. Justify your answer.

(a) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are diagonal matrices of the same size, then $\boldsymbol{AB} = \boldsymbol{BA}$.*

(b) *If $\boldsymbol{A}$ is a square matrix, then $\frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^T)$ is symmetric.*

(c) *For all matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, $(\boldsymbol{A} + \boldsymbol{B})^2 = \boldsymbol{A}^2 + \boldsymbol{B}^2 + 2\boldsymbol{AB}$.*

(d) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are symmetric matrices for the same size, then $\boldsymbol{A} - \boldsymbol{B}$ is symmetric.*

(e) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are symmetric matrices for the same size, then $\boldsymbol{AB}$ is symmetric.*

(f) *If $\boldsymbol{A}$ is a square matrix such that $\boldsymbol{A}^2 = \boldsymbol{0}$, then $\boldsymbol{A} = \boldsymbol{0}$.*

(g) *If $\boldsymbol{A}$ is a matrix such that $\boldsymbol{AA}^T = \boldsymbol{0}$, then $\boldsymbol{A} = \boldsymbol{0}$.*

(h) *There exists a real matrix $\boldsymbol{A}$, such that $\boldsymbol{A}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$.*

Solution.

(a) True. Let $\boldsymbol{A} = (a_{ij})_{n\times n}$ and $\boldsymbol{B} = (b_{ij})_{n\times n}$. Since $a_{ij} = b_{ij} = 0$ when $i \neq j$, the (i,j)-entry of $\boldsymbol{AB}$ is equal to

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \begin{cases} a_{ii}b_{ii} & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Likewise, the (i,j)-entry of $\boldsymbol{BA}$ is equal to

$$b_{i1}a_{1j} + b_{i2}a_{2j} + \cdots + b_{in}a_{nj} = \begin{cases} b_{ii}a_{ii} & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Thus, $\boldsymbol{AB} = \boldsymbol{BA}$.

(b) True. $\left[\frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^T)\right]^T = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^T)^T = \frac{1}{2}(\boldsymbol{A}^T + \boldsymbol{A})$.

(c) False. Choose any two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ which satisfy $\boldsymbol{AB} \neq \boldsymbol{BA}$. We will find that $(\boldsymbol{A}+\boldsymbol{B})^2 \neq \boldsymbol{A}^2 + \boldsymbol{B}^2 + 2\boldsymbol{AB}$.

(d) True. $(\boldsymbol{A} - \boldsymbol{B})^T = \boldsymbol{A}^T - \boldsymbol{B}^T = \boldsymbol{A} - \boldsymbol{B}$

$\square$

Solution (Cont.)

(e) False. Choose any two symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ which satisfy $\boldsymbol{AB} \neq \boldsymbol{BA}$. We will find that $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T = \boldsymbol{BA} \neq \boldsymbol{AB}$. For example:

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

(f) False. Example: $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

(g) True. Let $\boldsymbol{A} = (a_{ij})_{n\times n}$, then for each $i \in \{1, 2, \ldots, n\}$, (i, i)-entry of $\boldsymbol{AA}^T$ is equal to

$$a_{i1}a_{i1} + a_{i2}a_{i2} + \cdots + a_{in}a_{in} = \sum_{k=1}^{n} a_{ik}^2.$$

So $\boldsymbol{AA}^T = 0$ implies that $a_{ik} = 0$ for all i and k, i.e. $\boldsymbol{A} = \boldsymbol{0}$.

(h) True. Example: $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ or $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

$\square$

Remark

Compare $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ with $i \in \mathbb{C}$.

Exercise (2.19)

Let $\boldsymbol{A} = \begin{pmatrix} 2 & -1 & -1 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix}$.

(a) *Verify that* $\boldsymbol{A}^2 - 6\boldsymbol{A} + 8\boldsymbol{I} = \boldsymbol{0}$.

(b) *Show that* $\boldsymbol{A}^{-1} = \frac{1}{8}(6\boldsymbol{I} - \boldsymbol{A})$ *without computing the inverse of* $\boldsymbol{A}$ *explicitly.*

Solution and Proof.

(a)

$$\boldsymbol{A}^2 = \begin{pmatrix} 4 & -6 & -6 \\ 0 & 10 & 6 \\ 0 & 6 & 10 \end{pmatrix}, -6\boldsymbol{A} = \begin{pmatrix} -12 & 6 & 6 \\ 0 & -18 & -6 \\ 0 & -6 & -18 \end{pmatrix}, 8\boldsymbol{I} = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{pmatrix}.$$

It is easy to be checked that $\boldsymbol{A}^2 - 6\boldsymbol{A} + 8\boldsymbol{I} = \boldsymbol{0}$.

(b) By (a), $\boldsymbol{A}^2 - 6\boldsymbol{A} + 8\boldsymbol{I} = \boldsymbol{0}$, we have

$$\boldsymbol{I} = \frac{1}{8}\left[6\boldsymbol{A} - \boldsymbol{A}^2\right] = \boldsymbol{A}\left[\frac{1}{8}(6\boldsymbol{I} - \boldsymbol{A})\right].$$

By definition, $\boldsymbol{A}^{-1}$ is $\frac{1}{8}(6\boldsymbol{I} - \boldsymbol{A})$.

$\square$

Exercise (2.9)

Suppose the homogeneous system $\boldsymbol{Ax} = \mathbf{0}$ has non-trivial solution. Show that the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ has either no solution or infinitely many solutions.

Proof.

If $\boldsymbol{Ax} = \boldsymbol{b}$ has a solution $\boldsymbol{x} = \boldsymbol{u}$, then $\boldsymbol{u} + \boldsymbol{v}$ is also a solution to $\boldsymbol{Ax} = \boldsymbol{b}$ for all solutions $\boldsymbol{x} = \boldsymbol{v}$ to $\boldsymbol{Ax} = \mathbf{0}$ since

$$\boldsymbol{A}(\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{Au} + \boldsymbol{Av} = \boldsymbol{b} + \mathbf{0} = \boldsymbol{b}.$$

Hence $\boldsymbol{Ax} = \boldsymbol{b}$ has either no solutions or infinitely many solutions. □

Remark

The structure of the solution set of inhomogeneous system $\boldsymbol{Ax} = \boldsymbol{b}$:

$$\text{Solution set} = \{\boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{v} \text{ is any solution to } \boldsymbol{Ax} = \mathbf{0}\},$$

where $\boldsymbol{u}$ is any specific solution to the linear system $\boldsymbol{Ax} = \boldsymbol{b}$.

2nd Method.

- $\boldsymbol{Ax} = \mathbf{0}$ has non-trivial solution, then in a REF of its augmented matrix, # variables $>$ # pivot columns;
- For $\boldsymbol{Ax} = \boldsymbol{b}$, if in its REF, # pivot columns changes, then the last column must be a pivot column, i.e. $\boldsymbol{Ax} = \boldsymbol{b}$ can not be solved;
- For $\boldsymbol{Ax} = \boldsymbol{b}$, if in its REF, # pivot columns does not change, then the last column is not a pivot column, i.e. $\boldsymbol{Ax} = \boldsymbol{b}$ can be solved; At this time, # variables $>$ # pivot columns, i.e. $\boldsymbol{Ax} = \boldsymbol{b}$ has infinite solutions;

$\square$

Exercise (2.20)

Let $\boldsymbol{A}$ be a square matrix.

(a) *Show that if $\boldsymbol{A}^2 = \boldsymbol{0}$, then $\boldsymbol{I} - \boldsymbol{A}$ is invertible and $(\boldsymbol{I} - \boldsymbol{A})^{-1} = \boldsymbol{I} + \boldsymbol{A}$.*

(b) *Show that if $\boldsymbol{A}^3 = \boldsymbol{0}$, then $\boldsymbol{I} - \boldsymbol{A}$ is invertible and $(\boldsymbol{I} - \boldsymbol{A})^{-1} = \boldsymbol{I} + \boldsymbol{A} + \boldsymbol{A}^2$.*

(c) *If $\boldsymbol{A}^n = \boldsymbol{0}$ for $n \geq 4$, is $\boldsymbol{I} - \boldsymbol{A}$ invertible?*

Recall

Based on distributive law, $(\boldsymbol{I} - \boldsymbol{A})(\boldsymbol{I} + \boldsymbol{A} + \cdots + \boldsymbol{A}^{n-1}) = \boldsymbol{I} - \boldsymbol{A}^n$ where $n \geq 2$ is an integer.

Proof and Solution.

(a) Since $(\boldsymbol{I} - \boldsymbol{A})(\boldsymbol{I} + \boldsymbol{A}) = \boldsymbol{I} - \boldsymbol{A}^2 = \boldsymbol{I}$ and $(\boldsymbol{I} + \boldsymbol{A})(\boldsymbol{I} - \boldsymbol{A}) = \boldsymbol{I} - \boldsymbol{A}^2 = \boldsymbol{I}$, we have that $\boldsymbol{I} - \boldsymbol{A}$ is invertible and its inverse is $\boldsymbol{I} + \boldsymbol{A}$.

(b) Since $(\boldsymbol{I} - \boldsymbol{A})(\boldsymbol{I} + \boldsymbol{A} + \boldsymbol{A}^2) = \boldsymbol{I} - \boldsymbol{A}^3 = \boldsymbol{I}$ and $(\boldsymbol{I} + \boldsymbol{A} + \boldsymbol{A}^2)(\boldsymbol{I} - \boldsymbol{A}) = \boldsymbol{I} - \boldsymbol{A}^3 = \boldsymbol{I}$, we have that $\boldsymbol{I} - \boldsymbol{A}$ is invertible and its inverse is $\boldsymbol{I} + \boldsymbol{A} + \boldsymbol{A}^2$.

(c) Yes. $\boldsymbol{I} - \boldsymbol{A}$ is invertible and its inverse is $\boldsymbol{I} + \boldsymbol{A} + \cdots + \boldsymbol{A}^n$.

$\square$

Exercise (2.21)

(a) *Give three examples of* 2×2 *matrices* $\boldsymbol{A}$ *such that* $\boldsymbol{A}^2 = \boldsymbol{A}$.

(b) *Let* $\boldsymbol{A}$ *be a square matrix such that* $\boldsymbol{A}^2 = \boldsymbol{A}$. *Show that* $\boldsymbol{I} + \boldsymbol{A}$ *is invertible and* $(\boldsymbol{I} + \boldsymbol{A})^{-1} = \frac{1}{2}(2\boldsymbol{I} - \boldsymbol{A})$.

(c) *Is* $\boldsymbol{I} - \boldsymbol{A}$ *always invertible? (Question 5 in Final of 2006–2007(I))*

Method

For (b) and (c), suppose we have $(\boldsymbol{I} + \boldsymbol{A})(a\boldsymbol{I} + b\boldsymbol{A}) = \boldsymbol{I}$, and then solve a and b.

Solution and Proof.

(a) $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

(b) It is easy to obtain $a = 1$ and $b = -\frac{1}{2}$. Thus, $(\boldsymbol{I} + \boldsymbol{A})^{-1} = \boldsymbol{A} - \frac{1}{2}\boldsymbol{A}$.

(c) No. Example: take $\boldsymbol{A} = \boldsymbol{I}$.

□

Exercise (2.22)

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be invertible matrices of the same size.

(a) *Give an example of 2×2 invertible matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ such that $\boldsymbol{A} \neq -\boldsymbol{B}$ and $\boldsymbol{A} + \boldsymbol{B}$ is not invertible.*

(b) *If $\boldsymbol{A} + \boldsymbol{B}$ is invertible, show that $\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1}$ is invertible and $(\boldsymbol{A} + \boldsymbol{B})^{-1} = \boldsymbol{A}^{-1}(\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1}\boldsymbol{B}^{-1}$.*

Solution and Proof.

(a) Let $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, then $\boldsymbol{A}$ and $\boldsymbol{B}$ are invertible, $\boldsymbol{A} + \boldsymbol{B} \neq 0$, and $\boldsymbol{A} + \boldsymbol{B}$ is not invertible.

(b)

$$\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1} = \boldsymbol{B}^{-1}(\boldsymbol{B}\boldsymbol{A}^{-1} + \boldsymbol{I}) = \boldsymbol{B}^{-1}(\boldsymbol{B} + \boldsymbol{A})\boldsymbol{A}^{-1}.$$

Hence, $(\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1} = \boldsymbol{A}(\boldsymbol{A} + \boldsymbol{B})^{-1}\boldsymbol{B}$, i.e.

$$\boldsymbol{A}^{-1}(\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1}\boldsymbol{B}^{-1} = (\boldsymbol{A} + \boldsymbol{B})^{-1}.$$

□

Exercise (Question 2 in Final of 2001–2002(II))

Let $\boldsymbol{A}$ be an $n \times n$ matrix. Then $\boldsymbol{A}^2 = \boldsymbol{A}$ iff $(\boldsymbol{I} - 2\boldsymbol{A})^{-1} = \boldsymbol{I} - 2\boldsymbol{A}$.

Proof.

$\boldsymbol{A}^2 = \boldsymbol{A}$ iff $\boldsymbol{I} - 4\boldsymbol{A} + 4\boldsymbol{A}^2 = \boldsymbol{I}$ iff $(\boldsymbol{I} - 2\boldsymbol{A})^2 = \boldsymbol{I}$. □

Exercise (Question 1(b) in Final of 2005–2006(I))

Let $\boldsymbol{A}$ be an $m \times n$ matrix and $\boldsymbol{B}$ be an $n \times m$ matrix with $n < m$. Show that $\boldsymbol{AB}$ is singular.

Exercise (Question 1)

Let $\boldsymbol{A}$ be an $n \times n$ matrix, and $\boldsymbol{J}$ be an $n \times n$ matrix in which the all entries are 1. In each row of $\boldsymbol{A}$, there are exactly two entries which are 1, and others are 0. Find all matrices $\boldsymbol{A}$ which satisfy $\boldsymbol{A}^2 + 2\boldsymbol{A} = 2\boldsymbol{J}$.

Solution.

1. $$\boldsymbol{A}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} = \begin{pmatrix}2\\ \vdots\\ 2\end{pmatrix}, \quad \boldsymbol{A}^2\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} = 2\boldsymbol{A}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} = \begin{pmatrix}4\\ \vdots\\ 4\end{pmatrix}, \quad 2\boldsymbol{J}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} = \begin{pmatrix}2n\\ \vdots\\ 2n\end{pmatrix}$$
   Hence $4 + 4 = 2n$, i.e. $n = 4$, $\boldsymbol{A}$ is a matrix of order 4.
2. If $\boldsymbol{B}$ satisfies $\boldsymbol{A}^2 + 2\boldsymbol{A} = 2\boldsymbol{J}$, so does $\boldsymbol{B}^T$.
3. The task left is simple.

$\square$

Exercise (Question 2)

Given an invertible matrix, how to compute its inverse.

Exercise (Question 3)

When a matrix A is not invertible, how to extend the definition of inverse for A.

Change log

Last modified: 16:05, February 2, 2011.

Schedule of Tutorial 3

- Any question about last tutorial
- Review concepts:
  - Definition of determinant;
  - Cofactor expansion;
  - Properties and computation of determinant.
- Tutorial: 2.24, 2.26, 2.32, 2.35, 2.36, 2.37
- Additional material:
  - Two other equivalent definitions of determinant.
  - Laplace formula and Binet-Cauchy formula.
  - Determinants of Vandermonde matrix and Hilbert matrix.
  - Determinant for block matrices.

First definition of Determinant (Laplace formula)

- Let $\boldsymbol{A} = (a_{ij})$ be an $n \times n$ matrix. Let $\boldsymbol{M}_{ij}$ be an matrix obtained from $\boldsymbol{A}$ by deleting the ith row and the jth column. Then the determinant of $\boldsymbol{A}$ is defined as

$$\det(\boldsymbol{A}) = \begin{cases} a_{11} & \text{if } n = 1 \\ a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n} & \text{if } n \geq 2 \end{cases}$$

  where

$$A_{ij} = (-1)^{i+j} \det(\boldsymbol{M}_{ij}),$$

  which is called the (i, j)-cofactor of $\boldsymbol{A}$.

- Let $\boldsymbol{A} = (a_{ij})$ be an $n \times n$ matrix. $\det(\boldsymbol{A})$ is usually written as

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

Second definition of Determinant (Leibniz formula)

- A permutation of a set S is a bijection from S to itself. If S is a finite set of n elements, then there are $n!$ permutations of S. We use S_n to denote the set of all permutations of $\{1, 2, \ldots, n\}$.
- In the following notation, one lists the elements of S in the first row, and for each one its image under the permutation below it in the second row:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix},$$

  this means that σ satisfies $\sigma(1) = 2, \sigma(2) = 5, \sigma(3) = 4, \sigma(4) = 3$ and $\sigma(5) = 1$.
- If $S = \{1, 2, \ldots, n\}$, the parity of a permutation σ of S can be defined as the parity of the number of inversions for σ, i.e., of pairs of elements x, y of S such that $x < y$ and $\sigma(x) > \sigma(y)$.
- The sign or signature of a permutation σ is denoted $\operatorname{sgn}(\sigma)$ and defined as $+1$ if the parity of σ is even and -1 otherwise.
- Define

$$\det(\boldsymbol{A}) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Third definition of Determinant (Axioms)

Let D be a function from the set of all $n \times n$ matrices to $\mathbb{R}$.

- We say that D is n-linear if for each i $(1 \leq i \leq n)$, D is a linear function of the ith row when the other $(n-1)$ rows are held fixed.
- We say that D is alternating if the following two conditions are satisfied:
  - $D(\boldsymbol{A}) = 0$ whenever two rows of $\boldsymbol{A}$ are equal;
  - If $\boldsymbol{A}'$ is a matrix obtained from $\boldsymbol{A}$ by interchanging two rows of $\boldsymbol{A}$, then $D(\boldsymbol{A}') = -D(\boldsymbol{A})$.
- We say that D is a determinant function if D is n-linear, alternating, and $D(\boldsymbol{I}_n) = 1$.
- Existence: Corollary, page 147 in Hoffman's "Linear Algebra".
- Uniqueness: Theorem 2, page 152 in Hoffman's "Linear Algebra".
- Notation: $\det$.

Equivalence of the three definitions

Def 3 $\Rightarrow$ Def 1 Theorem 1, page 146 in Hoffman's "Linear Algebra";

Def 1 $\Rightarrow$ Def 3 Trivial;

Def 2 $\Rightarrow$ Def 1 Section 5.7, page 173–180 in Hoffman's "Linear Algebra", or 2.3 节，许以超，"线性代数与矩阵论";
Moreover, we will get Laplace Expansions (Ref Example 13, page 179 in Hoffman's "Linear Algebra", or 定理 2.3.3，许以超，"线性代数与矩阵论");

Def 1 $\Rightarrow$ Def 2 Mathematical Induction.

Properties of Determinants

- $\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$, and $\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + bfg + cdh - ceg - afh - bdi$.
- Let $\boldsymbol{A}$ be a square matrix. We can compute $\det(\boldsymbol{A})$ by performing cofactor expansion along any row or any column of $\boldsymbol{A}$:

$$\begin{aligned} \det(\boldsymbol{A}) &= a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} && \text{along } i\text{th row} \\ &= a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} && \text{along } j\text{th column} \end{aligned}$$

- Let $\boldsymbol{A}$ be a square matrix, then $\det(\boldsymbol{A}) = \det(\boldsymbol{A}^T)$; (By the last statement)
- Let $\boldsymbol{A}$ be a triangular (hence square) matrix, then $\det A$ is the product of its diagonal entries; (By induction and cofactor expansion)
- The determinant of a square matrix with two identical rows (columns) is zero; (By induction and cofactor expansion)
- $\det\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{C} \end{pmatrix} = \det(\boldsymbol{A})\det(\boldsymbol{C})$, where $\boldsymbol{A}$ and $\boldsymbol{C}$ are $m \times m$ and $n \times n$ square matrices, respectively. (See Determinant for block matrices.)

Properties of Determinants (Cont.)

- Let $\boldsymbol{A}$ be a square matrix.
  - If $\boldsymbol{E}$ is an elementary matrix of the same size as $\boldsymbol{A}$, then $\det(\boldsymbol{EA}) = \det(\boldsymbol{E})\det(\boldsymbol{A})$;
  - If $\boldsymbol{B}$ is obtained from $\boldsymbol{A}$ by multiplying one row of $\boldsymbol{A}$ by a constant k, then $\det(\boldsymbol{B}) = k\det(\boldsymbol{A})$;
  - If $\boldsymbol{B}$ is obtained from $\boldsymbol{A}$ by interchanging two rows of $\boldsymbol{A}$, then $\det(\boldsymbol{B}) = -\det(\boldsymbol{A})$;
  - If $\boldsymbol{B}$ is obtained from $\boldsymbol{A}$ by adding a multiple of one row of $\boldsymbol{A}$ to another row, then $\det(\boldsymbol{B}) = \det(\boldsymbol{A})$.
- Let $\boldsymbol{A}$ be a square matrix. Then $\boldsymbol{A}$ is invertible if and only if $\det(\boldsymbol{A}) \neq 0$.
- Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two square matrices of order n and c is a scalar. Then:
  - $\det(c\boldsymbol{A}) = c^n \det(\boldsymbol{A})$;
  - $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$;
  - If $\boldsymbol{A}$ is invertible, then $\det(\boldsymbol{A}^{-1}) = \frac{1}{\det(\boldsymbol{A})}$.

Additional Properties of Determinants

- Let $\boldsymbol{A} = (a_{ij})$ be an $m \times n$ matrix. For $1 \leq i_1 < i_2 < \cdots < i_r \leq m$, $1 \leq j_1 < j_2 < \cdots < j_s \leq n$, let

$$\boldsymbol{A}\binom{i_1\, i_2 \cdots i_r}{j_1 j_2 \cdots j_s} = \begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \cdots & a_{i_1 j_s} \\ a_{i_2 j_1} & a_{i_2 j_2} & \cdots & a_{i_2 j_s} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_r j_1} & a_{i_r j_2} & \cdots & a_{i_r j_s} \end{pmatrix}$$

- Laplace formula (Section 5.7 in Hoffman's "Linear Algebra", or 2.3 节，许以超，“线性代数与矩阵轮”): For $1 \leq i_1 < i_2 < \cdots < i_r \leq n$,

$$\det(\boldsymbol{A}) = \sum_{1 \leq j_1 < j_2 < \cdots < j_r \leq n} \det \boldsymbol{A}\binom{i_1 \cdots i_r}{j_1 \cdots j_r} \operatorname{sgn}\binom{i_1\, i_2 \cdots i_n}{j_1 j_2 \cdots j_n} \det \boldsymbol{A}\binom{i_{r+1} \cdots i_n}{j_{r+1} \cdots j_n},$$

  where $i_1 i_2 \cdots i_n$ and $j_1 j_2 \cdots j_n$ are permutations of $1, 2, \ldots, n$, and $1 \leq i_{r+1} < \cdots i_n \leq n$, $1 \leq j_{r+1} < \cdots j_n \leq n$.

- Binet-Cauchy formula (2.3 节，许以超，“线性代数与矩阵轮”): Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be $m \times n$ and $n \times m$ matrices, respectively. Then:

$$\det(\boldsymbol{AB}) = \begin{cases} 0 & \text{if } m > n \\ \det(\boldsymbol{A}) \det(\boldsymbol{B}) & \text{if } m = n \\ \sum_{1 \leq j_1 < \cdots < j_m \leq n} \det \boldsymbol{A}\binom{1 \cdots m}{j_1 \cdots j_m} \det \boldsymbol{B}\binom{j_1 \cdots j_m}{1 \cdots m} & \text{if } m < n \end{cases}$$

Exercise (2.24)

Consider the population of certain endangered species of wild animals: On the average, each adult will give birth to one baby each year; 50% of the new born babies will survive the first year; 60% of the one-year-old cubs will survive the second year and become adults; and 70% of the adults will survive each year.

Define $\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 1 \\ 0.5 & 0 & 0 \\ 0 & 0.6 & 0.7 \end{pmatrix}$. *Let* x_0, y_0 *and* z_0 *be the numbers of babies, one-year-old cubs and adults, respectively, at the end of a particular year.*

(a) *Let* $\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \boldsymbol{A} \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$. *What information do the numbers* x_1, y_1 *and* z_1 *give us?*

(b) *Let* $\begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} = \boldsymbol{A}^n \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$, *where* n *is a positive number. Interpret the numbers* x_n, y_n *and* z_n.

(c) *Suppose initially,* $x_0 = 0$, $y_0 = 0$, $z_0 = 100$. *What is the total population three years later?*

Solution.

In each year, there are three generations: babies, one-year-old cubs and adults.

(a)
- $x_1 = z_0$ is the number of babies at the end of the 1st year;
- $y_1 = 0.5x_0$ is the number of one-year-old cubs at the end of the 1st year;
- $z_1 = 0.6y_0 + 0.7z_0$ is the number of adults at the end of the 1st year.

(b)
- $x_2 = z_1$ is the number of babies at the end of the 2nd year;
- $y_2 = 0.5x_1$ is the number of one-year-old cubs at the end of the 2nd year;
- $z_2 = 0.6y_1 + 0.7z_1$ is the number of adults at the end of the 2nd year.

- $x_3 = z_2$ is the number of babies at the end of the 3rd year;
- $y_3 = 0.5x_2$ is the number of one-year-old cubs at the end of the 3rd year;
- $z_3 = 0.6y_2 + 0.7z_2$ is the number of adults at the end of the 3rd year.

Inductively, we will obtain: x_n, y_n and z_n are the numbers of babies, one-year-old cubs and adults, respectively, at the end of the nth year.

(c) Based on part (b), we have

$$\begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix} = \boldsymbol{A}^3 \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0.5 & 0 & 0 \\ 0 & 0.6 & 0.7 \end{pmatrix}^3 \begin{pmatrix} 0 \\ 0 \\ 100 \end{pmatrix} = \begin{pmatrix} 49 \\ 35 \\ 64.3 \end{pmatrix}.$$

Thus the total population three years later is 148 ($x_3 + y_3 + z_3 = 148.3 \doteq 148$).

$\square$

Exercise (2.26)

Let $\boldsymbol{A}$ be the 4×4 matrix obtained from $\boldsymbol{I}$ by the following sequence of elementary row operations:

$$\boldsymbol{I} \xrightarrow{\frac{1}{2}R_2} \xrightarrow{R_1 - R_2} \xrightarrow{R_2 \leftrightarrow R_4} \xrightarrow{R_3 + 3R_1} \boldsymbol{A}$$

(a) *Write $\boldsymbol{A}$ as a product of four elementary matrices.*

(b) *Find $\boldsymbol{A}^{-1}$ as a product of four elementary matrices.*

Solution of (a).

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{I}. \qquad \square$$

Solution of (b).

Since $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$, we have

$$
\begin{aligned}
\boldsymbol{A}^{-1} &= \left\{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{I} \right\}^{-1} \\
&= \boldsymbol{I}^{-1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1} \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
\end{aligned}
$$

$\square$

Exercise (2.32)

Solve the matrix equation $\begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 3 & 2 \end{pmatrix} \boldsymbol{X} = \begin{pmatrix} 2 & 3 & 4 & 1 \\ 1 & 0 & 3 & 7 \\ 2 & 1 & 1 & 2 \end{pmatrix}$.

Method

If $\begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 3 & 2 \end{pmatrix}$ is invertible, then $\boldsymbol{X}$ can be found easily.

Solution.

$$\left(\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 3 & 2 & 0 & 0 & 1 \end{array}\right) \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 4/7 & -1/7 & -1/7 \\ 0 & 1 & 0 & -2/7 & -3/7 & 4/7 \\ 0 & 0 & 1 & 1/7 & 5/7 & -2/7 \end{array}\right)$$

Thus $\begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 3 & 2 \end{pmatrix}$ is invertible and the inverse is $\frac{1}{7}\begin{pmatrix} 4 & -1 & -1 \\ -2 & -3 & 4 \\ 1 & 5 & -2 \end{pmatrix}$. So

$$\boldsymbol{X} = \frac{1}{7}\begin{pmatrix} 4 & -1 & -1 \\ -2 & -3 & 4 \\ 1 & 5 & -2 \end{pmatrix}\begin{pmatrix} 2 & 3 & 4 & 1 \\ 1 & 0 & 3 & 7 \\ 2 & 1 & 1 & 2 \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 5 & 11 & 12 & -5 \\ 1 & -2 & -13 & -15 \\ 3 & 1 & 17 & 32 \end{pmatrix}.$$

$\square$

Exercise (2.35)

(a) *Determine the values of* a, b *and* c *so that the homogeneous system*

$$\begin{cases} x + y + z = 0 \\ ax + by + cz = 0 \\ a^2x + b^2y + c^2z = 0 \end{cases}$$

has non-trivial solution.

(b) *Write down the conditions so that the matrix* $\begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{pmatrix}$ *is invertible.*

Solution.

(a) By Gaussian elimination, we have

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ a & b & c & 0 \\ a^2 & b^2 & c^2 & 0 \end{pmatrix} \xrightarrow[R_3 - a^2 R_1]{R_2 - aR_1} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & b-a & c-a & 0 \\ 0 & b^2-a^2 & c^2-a^2 & 0 \end{pmatrix}$$
$$\xrightarrow{R_3 - (a+b)R_2} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & b-a & c-a & 0 \\ 0 & 0 & (c-a)(c-b) & 0 \end{pmatrix}.$$

The homogeneous linear system has non-trivial solution if and only if # pivot points $<$ # variables. Here the necessary and sufficient condition is $b - a = 0$ or $(c-a)(c-b) = 0$, that is, $a = b$ or $a = c$ or $b = c$.

(b) $\begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{pmatrix}$ is invertible iff the homogeneous system has trivial solution, so by part (a), that is $a \neq b$, $a \neq c$ and $b \neq c$.

□

Remark

As we known, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ has non-trivial solution iff $\boldsymbol{A}$ is not invertible iff $\det(\boldsymbol{A}) = 0$. Hence, we may solve this question by computing $\det(\boldsymbol{A})$ directly. (In this question, $\det(\boldsymbol{A}) = (a-b)(a-c)(b-c)$.)

Exercise (2.36)

Let $\boldsymbol{A}$ be an $m \times n$ matrix and $\boldsymbol{B}$ an $n \times m$ matrix.

(a) *Suppose $\boldsymbol{A}$ is row equivalent to the following matrix: $\begin{pmatrix} \boldsymbol{R} \\ 0 \cdots 0 \end{pmatrix}$, where the last row is a zero and $\boldsymbol{R}$ is an $(m-1) \times n$ matrix. Show that $\boldsymbol{AB}$ is singular.*

(b) *If $m > n$, can $\boldsymbol{AB}$ be invertible? Justify your answer.*

(c) *When $m = 2$ and $n = 3$, give an example of $\boldsymbol{A}$ and $\boldsymbol{B}$ such that $\boldsymbol{AB}$ is invertible.*

Proof and Solution.

(a) Since $\boldsymbol{A}$ is row equivalent to $\begin{pmatrix} \boldsymbol{R} \\ 0 \cdots 0 \end{pmatrix}$, there exist some elementary matrices $\boldsymbol{E}_1, \ldots, \boldsymbol{E}_k$, such that $\boldsymbol{A} = \boldsymbol{E}_k \cdots \boldsymbol{E}_1 \begin{pmatrix} \boldsymbol{R} \\ 0 \cdots 0 \end{pmatrix}$. Hence $\boldsymbol{AB} = \boldsymbol{E}_k \cdots \boldsymbol{E}_1 \begin{pmatrix} \boldsymbol{RB} \\ 0 \cdots 0 \end{pmatrix}$, and $\boldsymbol{AB}$ can not be row equivalent to the identity matrix, i.e. $\boldsymbol{AB}$ is singular.

(b) Since a REF of $\boldsymbol{A}$ can have at most n non-zero rows and $m > n$, a row-echelon form of $\boldsymbol{A}$ must have a zero row. By part (a), $\boldsymbol{AB}$ cannot be invertible.

(c) For example, let $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ and $\boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$, then $\boldsymbol{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is invertible. (Here $\boldsymbol{BA} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ is not invertible.)

Exercise (2.37)

Determine which of the following statements are true. Justify your answer.

(a) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are invertible matrices of the same size, then $\boldsymbol{A}+\boldsymbol{B}$ is also invertible.*

(b) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are invertible matrices of the same size, then $\boldsymbol{AB}$ is also invertible.*

(c) *If $\boldsymbol{AB}$ is invertible where $\boldsymbol{A}$ and $\boldsymbol{B}$ are square matrices of the same size, then both $\boldsymbol{A}$ and $\boldsymbol{B}$ are invertible.*

Solution.

(a) False. For example, let $\boldsymbol{A}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ and $\boldsymbol{B}=\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$.

(b) True. See Theorem 2.3.10.

(c) True. Let $\boldsymbol{C}$ be the inverse of $\boldsymbol{AB}$. Then $\boldsymbol{A}(\boldsymbol{BC})=(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{I}$ which implies that $\boldsymbol{A}$ is invertible.
Likewise, $(\boldsymbol{CA})\boldsymbol{B}=\boldsymbol{C}(\boldsymbol{AB})=\boldsymbol{I}$ which implies that $\boldsymbol{B}$ is invertible.

□

Remark

For part (c), since $\boldsymbol{AB}$ is invertible, $\det(\boldsymbol{AB})\neq 0$. Thus $\det(\boldsymbol{A})\det(\boldsymbol{B})\neq 0$, and hence $\det(\boldsymbol{A})\neq 0$ and $\det(\boldsymbol{B})\neq 0$. Therefore $\boldsymbol{A}$ and $\boldsymbol{B}$ are invertible.

Exercise (Extra Question 1)

Find the determinant of Vandermonde[12] *matrix* $\boldsymbol{V}_n = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{pmatrix}$.

Solution.

For the case $n = 1$, it is trivial; so we focus on the case $n > 1$.
Performing $R_n - x_1 R_{n-1},\ R_{n-1} - x_1 R_{n-2},\ \cdots,\ R_2 - x_1 R_1$, we will get:

$$\det(\boldsymbol{V}_n) = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & x_2 - x_1 & x_3 - x_1 & \cdots & x_n - x_1 \\ 0 & x_2(x_2 - x_1) & x_3(x_3 - x_1) & \cdots & x_n(x_n - x_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & x_2^{n-2}(x_2 - x_1) & x_3^{n-2}(x_3 - x_1) & \cdots & x_n^{n-2}(x_n - x_1) \end{pmatrix}$$

$$= (-1)^{1+1} \det \begin{pmatrix} x_2 - x_1 & x_3 - x_1 & \cdots & x_n - x_1 \\ x_2(x_2 - x_1) & x_3(x_3 - x_1) & \cdots & x_n(x_n - x_1) \\ \vdots & \vdots & \ddots & \vdots \\ x_2^{n-2}(x_2 - x_1) & x_3^{n-2}(x_3 - x_1) & \cdots & x_n^{n-2}(x_n - x_1) \end{pmatrix}$$

□

[12] Alexandre-Théophile Vandermonde (February 28, 1735–January 1, 1796), a French musician and chemist.

Solution (Cont.)

$$\det(\boldsymbol{V}_n) = (x_2 - x_1)(x_3 - x_1)\cdots(x_n - x_1)\det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_2 & x_3 & \cdots & x_n \\ x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_2^{n-2} & x_3^{n-2} & \cdots & x_n^{n-2} \end{pmatrix}$$

$$= \prod_{j=2}^{n}(x_j - x_1)\prod_{j=3}^{n}(x_j - x_2)\det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_3 & x_4 & \cdots & x_n \\ x_3^2 & x_4^2 & \cdots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_3^{n-3} & x_4^{n-3} & \cdots & x_n^{n-3} \end{pmatrix}$$

Hence, by finite induction steps, we will obtain

$$\det(\boldsymbol{V}_n) = \prod_{i=1}^{n}\left[\prod_{j=i+1}^{n}(x_j - x_i)\right] = \prod_{1\le i<j\le n}(x_j - x_i).$$

□

Remark

$\boldsymbol{V}_n$ is invertible iff $x_i \neq x_j$ for all $i \neq j$.

Exercise (Extra Question 2)

Find the determinant of Hilbert[13] *matrix* $\boldsymbol{H}_n = \begin{pmatrix} \frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_n} \\ \frac{1}{a_2+b_1} & \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{a_n+b_1} & \frac{1}{a_n+b_2} & \cdots & \frac{1}{a_n+b_n} \end{pmatrix}$.

Solution.
We focus on the case $n > 1$.
Performing $R_n - R_1,\ R_{n-1} - R_1,\ \cdots,\ R_2 - R_1$, since

$$\frac{1}{a_i + b_j} - \frac{1}{a_1 + b_j} = \frac{a_1 - a_i}{(a_i + b_j)(a_1 + b_j)},$$

we will get:

$$\det(\boldsymbol{H}_n) = \frac{(a_1 - a_2)(a_1 - a_3)\cdots(a_1 - a_n)}{\prod_{j=1}^{n}(a_1 + b_j)} \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \frac{1}{a_2+b_1} & \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{a_n+b_1} & \frac{1}{a_n+b_2} & \cdots & \frac{1}{a_n+b_n} \end{pmatrix}.$$

□

[13] David Hilbert (January 23, 1862–February 14, 1943), a German mathematician.

Solution (Cont.)

Performing $C_n - C_1,\ C_{n-1} - C_1,\ \cdots,\ C_2 - C_1$, since

$$\frac{1}{a_i + b_1} - \frac{1}{a_i + b_j} = \frac{b_j - b_1}{(a_i + b_1)(a_i + b_j)},$$

we will get:

$$\det(\boldsymbol{H}_n) = \frac{\prod_{i=2}^{n}(a_1 - a_i)}{\prod_{j=1}^{n}(a_1 + b_j)} \frac{(b_1 - b_2)\cdots(b_1 - b_n)}{\prod_{j=2}^{n}(a_j + b_1)} \det\begin{pmatrix} \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_n} \\ \vdots & \ddots & \vdots \\ \frac{1}{a_n+b_2} & \cdots & \frac{1}{a_n+b_n} \end{pmatrix}.$$

Hence, by finite induction steps, we will obtain

$$\det(\boldsymbol{H}_n) = \frac{\prod_{1\le i<j\le n}(a_j - a_i)\prod_{1\le i<j\le n}(b_j - b_i)}{\prod_{i=1}^{n}\prod_{j=1}^{n}(a_i + b_j)}.$$

□

Exercise (Extra Question 3)

Let $G = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$, *then* $\det(G) = \det(A)\det(C)$ $(*)$, *where* A *and* C *are* $m \times m$ *and* $n \times n$ *square matrices, respectively.*

Proof.

Apply mathematical induction on m.

1. If $m = 1$, it is exactly cofactor expansion;
2. (i) For any $m \in \mathbb{N}$, assume the Equation $(*)$ holds, then we want to prove the Equation $(*)$ holds for $m + 1$:

   (ii) Performing cofactor expansion along first column, we will get:

   $$\det(G) = a_{11} G_{11} + a_{21} G_{21} + \cdots + a_{m1} G_{m1},$$

   where $G_{i1} = (-1)^{i+1} \det(M_{i1})$, where M_{i1} is the matrix obtained from G by deleting the ith row and the 1th column.

   (iii) Since $M_{i1} = \begin{pmatrix} N_{i1} & B \\ 0 & C \end{pmatrix}$, where N_{i1} is the matrix obtained from A by deleting the ith row and 1th column.

   (iv) By assumption, we have

   $$G_{i1} = (-1)^{i+1} \det \begin{pmatrix} N_{i1} & B \\ 0 & C \end{pmatrix} = (-1)^{i+1} \det(N_{i1}) \det(C) = A_{i1} \det(C),$$

   where A_{i1} is the $(i, 1)$-cofactor of A.

   (v) Hence, we have

   $$\det(G) = a_{11} A_{11} \det(C) + a_{21} A_{21} \det(C) + \cdots + a_{m1} A_{m1} \det(C) = \det(A)\det(C).$$

$\square$

Change log

Last modified: 15:01, February 12, 2011.

Schedule of Tutorial 4

- Any question about last tutorial
- Review concepts:
  - Matrices:
    - Adjoint: definition, properties;
    - Cramer's rule.
  - Vector spaces:
    - n-vector, Euclidean n-space;
    - Set notations for subsets of $\mathbb{R}^n$: implicit form and explicit form.
- Tutorial: 2.40, 2.48, 2.49, 3.4, Q3 in Mid-term Test of 2007–2008, Q4 in Mid-term Test of 2008–2009
- Additional material:
  - Question 2.46;
  - Question 3(b) in Mid-term Test of 2009–2010;
  - Proof for $\mathrm{adj}(\boldsymbol{AB}) = \mathrm{adj}(\boldsymbol{B})\,\mathrm{adj}(\boldsymbol{A})$;
  - Abstract definition of vector space.

(Classical) Adjoint

Let $\boldsymbol{A}$ be a square matrix of order $n > 1$.

- The (classical) adjoint of $\boldsymbol{A}$ is the $n \times n$ matrix

$$\operatorname{adj}(\boldsymbol{A}) = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix},$$

  where A_{ij} is the (i,j)-cofactor of $\boldsymbol{A}$.
- $\boldsymbol{A} \operatorname{adj}(\boldsymbol{A}) = \operatorname{adj}(\boldsymbol{A})\boldsymbol{A} = \det(\boldsymbol{A})\boldsymbol{I}_n$, no matter whether $\boldsymbol{A}$ is invertible;
-
  - If $\boldsymbol{A}$ is invertible, then $\boldsymbol{A}^{-1} = \frac{1}{\det(\boldsymbol{A})} \operatorname{adj}(\boldsymbol{A})$;
  - If $\boldsymbol{A}$ is invertible, then $\operatorname{adj}(\boldsymbol{A})$ is also invertible, and its inverse is $\frac{1}{\det(\boldsymbol{A})}\boldsymbol{A}$; (See Question 2.48.)
  - If $\boldsymbol{A}$ is invertible, then $\operatorname{adj}(\boldsymbol{A}^{-1}) = [\operatorname{adj}(\boldsymbol{A})]^{-1} = \frac{1}{\det(\boldsymbol{A})}\boldsymbol{A}$;
  - If $\operatorname{adj}(\boldsymbol{A})$ is invertible, then $\boldsymbol{A}$ is also invertible. (See Question 3(b) in Mid-term Test of 2009–2010.)
- $\operatorname{adj}(\boldsymbol{A}^T) = \operatorname{adj}(\boldsymbol{A})^T$; (By definition.)
- $\operatorname{adj}(c\boldsymbol{A}) = c^{n-1} \operatorname{adj}(\boldsymbol{A})$; (By definition.)
-
  - $\operatorname{adj}(\boldsymbol{AB}) = \operatorname{adj}(\boldsymbol{B}) \operatorname{adj}(\boldsymbol{A})$; (Using perturbation method.)
  - $\det(\operatorname{adj}(\boldsymbol{A})) = [\det(\boldsymbol{A})]^{n-1}$; (Using perturbation method.)
  - $\operatorname{adj}(\operatorname{adj}(\boldsymbol{A})) = [\det(\boldsymbol{A})]^{n-2}\boldsymbol{A}$. (Using perturbation method.)

Cramer's Rule

Let $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ be a linear system where $\boldsymbol{A}$ is an $n \times n$ matrix. Let $\boldsymbol{A}_i$ be the matrix obtained from $\boldsymbol{A}$ by replacing the ith column of $\boldsymbol{A}$ by $\boldsymbol{b}$. If $\boldsymbol{A}$ is invertible, then the system has only solution

$$\boldsymbol{x} = \frac{1}{\det(\boldsymbol{A})} \begin{pmatrix} \det(\boldsymbol{A}_1) \\ \det(\boldsymbol{A}_2) \\ \vdots \\ \det(\boldsymbol{A}_n) \end{pmatrix}.$$

n-vector and Euclidean n-space

- An n-vector has the form $(u_1, u_2, \ldots, u_i, \ldots, u_n)$, where u_1, u_2, …, u_n are real numbers, and u_i is the ith coordinate.
- Let $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ be two n-vectors.
  - We say that $\boldsymbol{u}$ and $\boldsymbol{v}$ are equal iff $u_i = v_i$ for all $i = 1, 2, \ldots, n$;
  - The addition $\boldsymbol{u} + \boldsymbol{v}$ is defined by $\boldsymbol{u} + \boldsymbol{v} = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$;
  - Let c be a real number. The scalar multiple $c\boldsymbol{u}$ is defined by $c\boldsymbol{u} = (cu_1, cu_2, \ldots, cu_n)$;
  - The n-vector $(0, 0, \ldots, 0)$ is called the zero vector and denote it by $\mathbf{0}$.
- We identify an n-vector $(u_1, u_2, \ldots, u_n)$ with a $1 \times n$ matrix $\begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix}$ or an $n \times 1$ matrix $\begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix}^T$.
- Let $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ be n-vectors and a, b real numbers. Then
  - (1) $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$;
  - (2) $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$;
  - (3) $\boldsymbol{u} + \mathbf{0} = \boldsymbol{u} = \mathbf{0} + \boldsymbol{u}$;
  - (4) $\boldsymbol{u} + (-\boldsymbol{u}) = \mathbf{0}$;
  - (5) $a(b\boldsymbol{u}) = (ab)\boldsymbol{u}$;
  - (6) $a(\boldsymbol{u} + \boldsymbol{v}) = a\boldsymbol{u} + a\boldsymbol{v}$;
  - (7) $(a + b)\boldsymbol{u} = a\boldsymbol{u} + b\boldsymbol{u}$;
  - (8) $1\boldsymbol{u} = \boldsymbol{u}$.
- The set of all n-vectors of real numbers space is called the Euclidean n-space and is denoted by $\mathbb{R}^n$.

Set notations for subsets of $\mathbb{R}^n$

- Set notation for subsets of $\mathbb{R}^n$:
  - Implicit form: $\{(u_1, u_2, \ldots, u_n) \mid$ conditions satisfied by $u_1, u_2, \ldots, u_n\}$;
  - Explicit form: $\{n$-vectors in terms of some parameters $\mid$ range of the parameters$\}$.
- Examples:
  - Lines in xy-plane: $\begin{cases} \text{Implicit form: } \{(x, y) \mid ax + by = c\} \\ \text{Explicit form: } \{(\text{general solution}) \mid 1 \text{ parameter}\} \end{cases}$
  - Planes in xyz-space: $\begin{cases} \text{Implicit form: } \{(x, y, z) \mid ax + by + cz = d\} \\ \text{Explicit form: } \{(\text{general solution}) \mid 2 \text{ parameters}\} \end{cases}$
  - Lines in xyz-space: $\begin{cases} \text{Implicit form: } \{(x, y, z) \mid \text{eqn of the line}\} \\ \text{Explicit form: } \{(\text{general solution}) \mid 1 \text{ parameter}\} \end{cases}$

Exercise (2.40)

Let $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & -1 \\ -2 & 1 & 0 & -2 \\ 0 & 0 & 2 & 1 \end{pmatrix}$, $\boldsymbol{C} = \begin{pmatrix} -1 & 3 & 4 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, $\boldsymbol{b} = \begin{pmatrix} 2 \\ 4 \\ 6 \\ 8 \end{pmatrix}$, $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$.

(a) *Solve the linear system* $\boldsymbol{Ax} = \boldsymbol{b}$.

(b) *Without computing the matrix* $\boldsymbol{AC}$, *explain why the homogeneous linear system* $\boldsymbol{ACx} = \boldsymbol{0}$ *has infinitely many solutions.*

Solution.

(a) $\left(\begin{array}{cccc|c} 1 & 0 & 2 & 0 & 2 \\ 0 & 1 & 3 & -1 & 4 \\ -2 & 1 & 0 & -2 & 6 \\ 0 & 0 & 2 & 1 & 8 \end{array}\right) \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -\frac{22}{3} \\ 0 & 1 & 0 & 0 & -\frac{34}{3} \\ 0 & 0 & 1 & 0 & \frac{14}{3} \\ 0 & 0 & 0 & 1 & -\frac{4}{3} \end{array}\right)$. So $x_1 = -\frac{22}{3}$, $x_2 = -\frac{34}{3}$, $x_3 = \frac{14}{3}$ and $x_4 = -\frac{4}{3}$.

(b)
1. Since $\boldsymbol{C}$ is an upper-triangular matrix, $\det(\boldsymbol{C})$ is the product of its diagonal entries, which is $-1 \times 0 \times 0 \times 1 = 0$.
2. Since $\det(\boldsymbol{AC}) = \det(\boldsymbol{A})\det(\boldsymbol{C}) = 0$, the homogeneous system $\boldsymbol{ACx} = \boldsymbol{0}$ has infinitely many solutions.

□

Exercise (2.48)

Let $\boldsymbol{A}$ be an $n \times n$ invertible matrix.

(a) *Show that* $\mathrm{adj}(\boldsymbol{A})$ *is invertible.*

(b) *Find* $\det(\mathrm{adj}(\boldsymbol{A}))$ *and* $\mathrm{adj}(\boldsymbol{A})^{-1}$.

(c) *If* $\det(\boldsymbol{A}) = 1$, *show that* $\mathrm{adj}(\mathrm{adj}(\boldsymbol{A})) = \boldsymbol{A}$.

Recall

Let $\boldsymbol{A}$ be a square matrix of order n.

1. $\boldsymbol{A}$ is invertible iff there exists a square matrix $\boldsymbol{B}$ of order n such that $\boldsymbol{AB} = \boldsymbol{I}$ and $\boldsymbol{BA} = \boldsymbol{I}$.
2. $\boldsymbol{A}$ is invertible iff there exists a square matrix $\boldsymbol{B}$ of order n such that $\boldsymbol{AB} = \boldsymbol{I}$.
3. $\boldsymbol{A}$ is invertible iff there exists a square matrix $\boldsymbol{B}$ of order n such that $\boldsymbol{BA} = \boldsymbol{I}$.

Proof and solution.

(a) Since $\mathrm{adj}(\boldsymbol{A})$ is a square matrix, $\boldsymbol{A}\,\mathrm{adj}(\boldsymbol{A}) = \det(\boldsymbol{A})\boldsymbol{I}_n$ and $\det(\boldsymbol{A}) \neq 0$, we have that $\mathrm{adj}(\boldsymbol{A})$ is invertible and its inverse is $\frac{1}{\det(\boldsymbol{A})}\boldsymbol{A}$.

□

Proof and Solution (Cont.)

(b) Since $\boldsymbol{A}\,\mathrm{adj}(\boldsymbol{A}) = \det(\boldsymbol{A})\boldsymbol{I}_n$, we have

$$\det(\boldsymbol{A})\det(\mathrm{adj}(\boldsymbol{A})) = \det(\boldsymbol{A}\,\mathrm{adj}(\boldsymbol{A})) = \det(\det(\boldsymbol{A})\boldsymbol{I}_n) = \det(\boldsymbol{A})^n.$$

Since $\det(\boldsymbol{A}) \neq 0$, we have $\det(\mathrm{adj}(\boldsymbol{A})) = \det(\boldsymbol{A})^{n-1}$.

(c)
1. For any square matrix $\boldsymbol{X}$, $\boldsymbol{X}\,\mathrm{adj}(\boldsymbol{X}) = \det(\boldsymbol{X})\boldsymbol{I}_n$.
2. Taking $\boldsymbol{X}$ to be $\mathrm{adj}(\boldsymbol{A})$ and by part (b), we will have

$$\mathrm{adj}(\boldsymbol{A})\,\mathrm{adj}(\mathrm{adj}(\boldsymbol{A})) = \det(\mathrm{adj}(\boldsymbol{A}))\boldsymbol{I}_n = \det(\boldsymbol{A})^n\boldsymbol{I}_n = \boldsymbol{I}_n.$$

3. We also have $\boldsymbol{A}\,\mathrm{adj}(\boldsymbol{A}) = \det(\boldsymbol{A})\boldsymbol{I}_n = \boldsymbol{I}_n$.
4. By definition, both $\boldsymbol{A}$ and $\mathrm{adj}(\mathrm{adj}(\boldsymbol{A}))$ are the inverse of $\boldsymbol{A}$, and hence they are the same.

$\square$

Exercise (2.49)

Determine which of the following statements are true. Justify your answer.

(a) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are square matrices of the same size, then* $\det(\boldsymbol{A}+\boldsymbol{B}) = \det(\boldsymbol{A}) + \det(\boldsymbol{B})$.

(b) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are square matrices of the same size, then* $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$.

(c) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are square matrices of the same size such that $\boldsymbol{A} = \boldsymbol{PBP}^{-1}$ for some invertible matrix $\boldsymbol{P}$, then* $\det(\boldsymbol{A}) = \det(\boldsymbol{B})$.

(d) *If $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are invertible matrices of the same size such that* $\det(\boldsymbol{A}) = \det(\boldsymbol{B})$, *then* $\det(\boldsymbol{A}+\boldsymbol{C}) = \det(\boldsymbol{B}+\boldsymbol{C})$.

Solution.

(a) False. For example, let $\boldsymbol{A} = \boldsymbol{I}_2$ and $\boldsymbol{B} = -\boldsymbol{I}_2$.

(b) True. See Theorem 2.5.27.

(c) True. Because $\det(\boldsymbol{A}) = \det(\boldsymbol{P})\det(\boldsymbol{B})\det(\boldsymbol{P}^{-1})$ and $\det(\boldsymbol{P})\det(\boldsymbol{P}^{-1}) = 1$.

(d) False. For example, let $\boldsymbol{A} = \boldsymbol{I}_2$ and $\boldsymbol{B} = \boldsymbol{C} = -\boldsymbol{I}_2$.

$\square$

Exercise (3.4)

Consider the following subsets of $\mathbb{R}^3$*:*

$$\begin{aligned}
A &= \textit{a line passes through the origin and } (9,9,9), \\
B &= \{(k,k,k) \mid k \in \mathbb{R}\}, \\
C &= \{(x_1,x_2,x_3) \mid x_1 = x_2 = x_3\}, \\
D &= \{(x,y,z) \mid 2x - y - z = 0\}, \\
E &= \{(a,b,c) \mid 2a - b - c = 0 \textit{ and } a + b + c = 0\}, \\
F &= \{(u,v,w) \mid 2u - v - w = 0 \textit{ and } 3u - 2v - w = 0\}.
\end{aligned}$$

Which of these subsets are the same?

Method

If we can express the sets in explicit form, then it is easy to compare them.

Solution.

- It is obvious that $A = B = C$;
- By solving the linear system, we have $F = C = B = A$;
- Since $D = \{(\frac{s+t}{2}, s, t) \mid s,t \in \mathbb{R}\}$ and $E = \{(0, s, -s) \mid s \in \mathbb{R}\}$, A, D and E are different with each other.

$\square$

Exercise (Question 3 in Mid-term Test of 2007–2008)

Consider the following subsets of $\mathbb{R}^3$*. (Note that vectors in* $\mathbb{R}^3$ *can be written in row or column form and regarded as the same.)*

$$S = \{(x, y, z) \mid 2x - 3y + z = 10 \text{ and } x - z = 5\}$$

$$T = \text{the solution set of the linear system } \begin{pmatrix} 2 & -3 & 1 \\ 1 & 0 & -1 \\ 3 & -3 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 10 \\ 5 \\ 15 \end{pmatrix}$$

$$U = \{(t + 2, t - 3, t - 3) \mid t \in \mathbb{R}\}$$

(i) *Determine whether the vector* $(3, -2, -2)$ *belongs to each of the three sets.*

(ii) *Describe the three sets geometrically (i.e. whether they represent points, lines, planes or others).*

(iii) *Which of the three sets are the same, if any? Justify your answers.*

Solution.

(i)

- Since $(3, -2, -2)$ satisfies $2x - 3y + z = 10$ and $x - z = 5$, $(3, -2, -2) \in S$;
- Since $(3, -2, -2)$ satisfies $\begin{pmatrix} 2 & -3 & 1 \\ 1 & 0 & -1 \\ 3 & -3 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 10 \\ 5 \\ 15 \end{pmatrix}$, $(3, -2, -2) \in T$;
- Taking t to be 1, we get $(t+2, t-3, t-3) = (3, -2, -2)$, hence $(3, -2, -2) \in U$.

(ii, iii)

- $$\left(\begin{array}{ccc|c} 2 & -3 & 1 & 10 \\ 1 & 0 & -1 & 5 \end{array}\right) \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 5 \\ 0 & 1 & -1 & 0 \end{array}\right).$$
  So $S = \{(s+5, s, s) \mid s \in \mathbb{R}\}$;
- $$\left(\begin{array}{ccc|c} 2 & -3 & 1 & 10 \\ 1 & 0 & -1 & 5 \\ 3 & -3 & 0 & 15 \end{array}\right) \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 5 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$
  So $T = \{(k+5, k, k) \mid k \in \mathbb{R}\}$;
- $U = \{(t+2, t-3, t-3) \mid t \in \mathbb{R}\} = \{(t+5, t, t) \mid t \in \mathbb{R}\}$.

Hence, all of them are same, and each of them represents a line in $\mathbb{R}^3$ since there is 1 free parameter.

$\square$

Exercise (Question 4 in Mid-term Test of 2008–2009)

Let P represent a plane in the xyz-space with equation $x - y + z = 1$ and A, B, C represent three different lines given by the following set notations:

$$A = \{(a, a, 1) \mid a \in \mathbb{R}\}, \quad B = \{(b, 0, 0) \mid b \in \mathbb{R}\}, \quad C = \{(c, 0, -c) \mid c \in \mathbb{R}\}.$$

(a) *Write down an explicit set notation that represents the plane P.*

(b) *Does any of the three lines above lie completely on the plane P? Briefly explain your answer.*

(c) *Find all the points of intersection of the line B with the plane P.*

(d) *Find the equation of another plane that is parallel to (but not overlapping) the plane P, and contains exactly one of the above three lines.*

(e) *Can you find a linear system whose solution set contains all the three lines? Justify your answer.*

Solution.

(a) By finding the general solution of the equation $x - y + z = 1$, we get the explicit form $\{(1 + s - t, s, t) \mid s, t \in \mathbb{R}\}$.

(b) A lies on the plane, as any point $(a, a, 1) \in A$ satisfies $x - y + z = 1$ $(a - a + 1 = 1)$.

(c) By substituting a point $(b, 0, 0) \in B$ into $x - y + z = 1$, we see that the only point in B that satisfies the equation is when $b = 1$.

(d) A plane that is parallel to P has the form $x - y + z = k$ $(k \neq 1)$. Such a plane will not intersection P, and so cannot contain lines A and B. Line C does not lie on P as none of the point $(c, 0, -c) \in C$ satisfies $x - y + z = 1$. Instead C lies on $x - y + z = 0$.

(e) No. The solution set must either be a point, a line or a plane. But there is no plane in xyz-space that contains all the three lines. (Take $(1, 1, 1) \in A$, $(1, 0, 0) \in B$, $(1, 0, 1) \in C$ and check that there is no equation $dx + ey + fz = g$ that is simultaneously satisfied by these three points.)

$\square$

Exercise (2.46)

Let $\boldsymbol{A} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ *where* $a, b, c, d, e, f, g, h, i$ *are either 0 or 1. Find the largest possible value and the smallest possible value of* $\det(\boldsymbol{A})$*.*

Solution.

$\det(\boldsymbol{A}) = aei + bfg + cdh - afh - bdi - ceg.$

- If all $a, b, c, d, e, f, g, h, i$ are 1, then $\det(\boldsymbol{A}) = 0$.
- Suppose at least one of $a, b, c, d, e, f, g, h, i$ is 0, say $a = 0$ (other cases are similar). Then $\det(\boldsymbol{A}) = bfg + cdh - bdi - ceg$. As b, c, d, e, f, g, h, i can only be 0 and 1, $-2 \leq \det(\boldsymbol{A}) \leq 2$.
- Note that $\begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{vmatrix} = 2$ and $\begin{vmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix} = -2$.
- The maximum possible value of $\det(\boldsymbol{A})$ is 2 and the minimum is -2.

$\square$

Exercise (Question 3(b) in Mid-term Test of 2009–2010)

If $\mathrm{adj}(\boldsymbol{A})$ *is invertible, then* $\boldsymbol{A}$ *is also invertible.*

Proof.

1. Assume that $\boldsymbol{A}$ is not invertible, then $\det(\boldsymbol{A}) = 0$;
2. Then $\boldsymbol{A}\,\mathrm{adj}(\boldsymbol{A}) = \det(\boldsymbol{A})\boldsymbol{I}_n = \boldsymbol{0}_{n\times n}$;
3. Since $\mathrm{adj}(\boldsymbol{A})$ is invertible, we have $\boldsymbol{A} = \boldsymbol{A}\,\mathrm{adj}(\boldsymbol{A})\,\mathrm{adj}(\boldsymbol{A})^{-1} = \boldsymbol{0}_{n\times n}$;
4. Hence $\mathrm{adj}(\boldsymbol{A}) = \boldsymbol{0}_{n\times n}$, which is a contradiction. Hence $\boldsymbol{A}$ is invertible.

$\square$

Exercise

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two square matrices of order $n > 1$. Then $\operatorname{adj}(\boldsymbol{AB}) = \operatorname{adj}(\boldsymbol{B})\operatorname{adj}(\boldsymbol{A})$.

Proof.

Assume $\boldsymbol{A}$ and $\boldsymbol{B}$ are invertible, then

$$\operatorname{adj}(\boldsymbol{AB}) = \det(\boldsymbol{AB})(\boldsymbol{AB})^{-1} = \det(\boldsymbol{B})\boldsymbol{B}^{-1}\det(\boldsymbol{A})\boldsymbol{A}^{-1} = \operatorname{adj}(\boldsymbol{B})\operatorname{adj}(\boldsymbol{A}).$$

Assume $\boldsymbol{A}$ is not invertible, and $\boldsymbol{B}$ is invertible.

(1) For $t > 0$, consider $\boldsymbol{A}_t = \boldsymbol{A} + t\boldsymbol{I}_n$. Then $\det(\boldsymbol{A}_t)$ is a polynomial of degree n, and hence $\det(\boldsymbol{A}_c) = 0$ has at most n solutions, say $t_1, t_2, \ldots, t_n$.

(2)
- If not all of $t_1, t_2, \ldots, t_n$ are 0, then we can find $\delta > 0$, such that $\delta < \min\{|t_k| \mid t_k \neq 0, 1 \leq k \leq n\}$. Then for any $t \in (0, \delta)$, $\det(\boldsymbol{A}_t) \neq 0$;
- If all of $t_1, t_2, \ldots, t_n$ are 0, then choosing arbitrary positive real number δ, we will have $\det(\boldsymbol{A}_t) \neq 0$ for any $t \in (0, \delta)$.

(3) Based on the discussion above, we can find $\delta > 0$, such that $\det(\boldsymbol{A}_t) \neq 0$ for any $t \in (0, \delta)$. Thus $\boldsymbol{A}_t$ is invertible for $t \in (0, \delta)$, and hence $\operatorname{adj}(\boldsymbol{A}_t\boldsymbol{B}) = \operatorname{adj}(\boldsymbol{B})\operatorname{adj}(\boldsymbol{A}_t)$.

□

Proof (Cont.)

(4) We use $l_{ij}(t)$ and $r_{ij}(t)$ denote $\mathrm{adj}(\boldsymbol{A}_t\boldsymbol{B})$'s and $\mathrm{adj}(\boldsymbol{B})\,\mathrm{adj}(\boldsymbol{A}_t)$'s (i,j)-entries, respectively.

(5) It is obvious that $l_{ij}(t)$ and $r_{ij}(t)$ are polynomials in term of t, and hence they are continuous. Since $l_{ij}(t) = r_{ij}(t)$ for any $t \in (0, \delta)$, we have

$$l_{ij}(0) = \lim_{t\to 0} l_{ij}(t) = \lim_{t\to 0} r_{ij}(t) = r_{ij}(t),$$

and hence $\lim_{t\to 0} \boldsymbol{A}_t = \boldsymbol{A}$.

(6) By the similar method, we have $\lim_{t\to 0} \mathrm{adj}(\boldsymbol{A}_t\boldsymbol{B}) = \mathrm{adj}(\boldsymbol{A}\boldsymbol{B})$, and hence

$$\mathrm{adj}(\boldsymbol{A}\boldsymbol{B}) = \lim_{t\to 0} \mathrm{adj}(\boldsymbol{A}_t\boldsymbol{B}) = \lim_{t\to 0} \mathrm{adj}(\boldsymbol{B})\,\mathrm{adj}(\boldsymbol{A}_t) = \mathrm{adj}(\boldsymbol{B})\,\mathrm{adj}(\boldsymbol{A}).$$

Assume $\boldsymbol{A}$ and $\boldsymbol{B}$ are not invertible. Then by the similar method, we also have $\mathrm{adj}(\boldsymbol{A}\boldsymbol{B}) = \mathrm{adj}(\boldsymbol{B})\,\mathrm{adj}(\boldsymbol{A})$. $\square$

Abstract definition of vector space

A vector space (or linear space) consists of the following:

1. a field $\mathbb{F}$ of scalars;
2. a set V of objects, called vectors;
3. an operation, called vector addition, which associated with each pair of vectors $\boldsymbol{u}$, $\boldsymbol{v}$ in V, called the sum of $\boldsymbol{u}$ and $\boldsymbol{v}$, in such a way that
   (1) addition is commutative, $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$;
   (2) addition is associated, $\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})=(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}$;
   (3) there is a unique vector $\mathbf{0}$ in V, called the zero vector, such that $\boldsymbol{u}+\mathbf{0}=\boldsymbol{u}$ for all $\boldsymbol{u}\in V$;
   (4) for each vector $\boldsymbol{u}$ in V there is a unique vector $-\boldsymbol{u}$ in V such that $\boldsymbol{u}+(-\boldsymbol{u})=\mathbf{0}$;
4. an operation, called scalar multiplication, which associates with each scalar c in $\mathbb{F}$ and a vector $\boldsymbol{u}$ in V a vector $c\boldsymbol{u}$ in V, called the product of c and $\boldsymbol{u}$, in such a way that
   (1) $1\boldsymbol{u}=\boldsymbol{u}$ for all $\boldsymbol{u}$ in V;
   (2) $(c_1c_2)\boldsymbol{u}=c_1(c_2\boldsymbol{u})$;
   (3) $c(\boldsymbol{u}+\boldsymbol{v})=c\boldsymbol{u}+c\boldsymbol{v}$;
   (4) $(c_1+c_2)\boldsymbol{u}=c_1\boldsymbol{u}+c_2\boldsymbol{u}$.

Change log

- Page 96: Change "$(t-2, t+3, t+3)$" to "$(t+2, t-3, t-3)$";
- Page 101: Add a proof for $\mathrm{adj}(\boldsymbol{AB}) = \mathrm{adj}(\boldsymbol{B})\,\mathrm{adj}(\boldsymbol{A})$.

Last modified: 00:30, February 19, 2011.

Information of Mid-Term Test

- Time: March 3rd, 18:00–19:00;
- Venue: MPSH1;
- Close book with 1 helpsheet;
- Consultation: March 2nd, 3rd
  - Office: S17-06-14.
  - Mobile: 9053-5550.
  - Email: xiangsun@nus.edu.sg.

Schedule of Tutorial 5

- Any question about last tutorial
- Review concepts: Vector spaces:
  - Linear combination, linear span;
  - Subspace;
  - Linear independence.
- Tutorial: 3.6, 3.12, 3.16, 3.18, 3.22, 3.23, 3.24
- Additional material: 3.10(a), 3.20, 3.21

Linear combination, linear span, and subspace

- $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are fixed vectors in $\mathbb{R}^n$, and $c_1, c_2, \ldots, c_k$ are real numbers. $c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k$ is called a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$.
- $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$: a (finite) subset of $\mathbb{R}^n$. The set of all linear combinations of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$

  $$\{c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k \mid c_1, c_2, \ldots, c_k \in \mathbb{R}\}$$

  is called the linear span of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$, or the linear span of S. Natation: $\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ or $\operatorname{span}(S)$.
- Let V be a subset of $\mathbb{R}^n$. V is called a subspace of $\mathbb{R}^n$ if there exists a set $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ of $\mathbb{R}^n$ such that $V = \operatorname{span}(S)$.
-
  - If V is a subspace of $\mathbb{R}^n$, then the zero vector $\mathbf{0} \in V$. (Hence a subspace can not be empty.)
  - Let V be a subspace of $\mathbb{R}^n$. If $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k \in V$, and $c_1, c_2, \ldots, c_k \in \mathbb{R}$, then $c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k \in V$.
- Exercise 3.21: Let V be a non-empty subset of $\mathbb{R}^n$. V is a subspace iff V satisfies the closed condition: for any $\boldsymbol{u}, \boldsymbol{v} \in V$ and $a, b \in \mathbb{R}$, $a\boldsymbol{u} + b\boldsymbol{v} \in V$.
-
  - $\{\mathbf{0}\}$, lines through the origin and $\mathbb{R}^2$ are all the subspaces of $\mathbb{R}^2$;
  - $\{\mathbf{0}\}$, lines through the origin, planes containing the origin and $\mathbb{R}^3$ are all the subspaces of $\mathbb{R}^3$.
-
  - The solution set of every homogeneous linear system is a subspace of $\mathbb{R}^n$;
  - The solution set of every inhomogeneous linear system is not a subspace of $\mathbb{R}^n$.

Methods for proving or disproving subspace

- We have FOUR methods for showing a subset $V \subset \mathbb{R}^n$ to be a subspace:
  - Express V as a linear span;
  - (By Exercise 3.21) show that V is non-empty and satisfies the closed condition;
  - Show that V is a solution set of some homogeneous linear system;
  - (For $\mathbb{R}^2$ and $\mathbb{R}^3$) show that V represents a line or plane through origin.

  First two methods are general, while the last two are available for some special cases.
- We have FIVE methods for showing a subset $V \subset \mathbb{R}^n$ to be not a subspace:
  - Show that zero vector in not in V;
  - Find $\boldsymbol{u}, \boldsymbol{v} \in V$, such that $\boldsymbol{u} + \boldsymbol{v} \notin V$;
  - Find $\boldsymbol{u} \in V$ and a scalar $c \in \mathbb{R}$, such that $c\boldsymbol{u} \notin V$;
  - Show that V is a solution set of some inhomogeneous linear system;
  - (For $\mathbb{R}^2$ and $\mathbb{R}^3$) show that V is not a line or plane through origin.

  First three methods are general, while the last two are available for some special cases.

Operations of subspaces

Let V and W be subspaces of $\mathbb{R}^n$.

- Define $V + W = \{\boldsymbol{v} + \boldsymbol{w} \mid \boldsymbol{v} \in V, \boldsymbol{w} \in W\}$, then $V + W$ is a subspace of $\mathbb{R}^n$. See Exercise 3.10(a).
- $V \cap W$ is a subspace of $\mathbb{R}^n$. See Exercise 3.22(a).
- $V \cup W$ is a subspace of $\mathbb{R}^n$ iff $V \subset W$ or $W \subset V$. See Exercise 3.22(c).
- Difference between $V + W$ and $V \cup W$: take V to be the x-axis, and W to be the y-axis in $\mathbb{R}^2$. It is obvious that V and W are subspaces in $\mathbb{R}^2$.

  By definition, we can see that $V + W = \mathbb{R}^2$: for any vector $\boldsymbol{u} \in \mathbb{R}^2$, we can write it as $\boldsymbol{u} = (x_1, y_1)$. Let $\boldsymbol{v} = x_1(1, 0)$, $\boldsymbol{w} = y_1(0, 1)$, then $\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{w}$. It is easy to see $\boldsymbol{v} \in V$ and $\boldsymbol{w} \in W$. Thus $\boldsymbol{u} \in V + W$.

  While $V \cup W$ is the union of the x-axis and the y-axis, which is not a subspace because $(1, 0)$ and $(0, 1)$ are in $V \cup W$, but $(1, 1)$ not.

Linear independence

- Problem 3.2.13: Suppose $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are vectors taken from $\mathbb{R}^n$. Show that if $\boldsymbol{u}_k$ is a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{k-1}$, then

$$\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{k-1}\} = \operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{k-1}, \boldsymbol{u}_k\}.$$

- Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\} \subset \mathbb{R}^n$.
  - S is called a linearly independent set and $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are said to be linearly independent if the vector equation

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k = \mathbf{0}$$

    has only trivial solution, where $c_1, c_2, \ldots, c_k$ are variables.
  - S is called a linearly dependent set and $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are said to be linearly dependent if the vector equation

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k = \mathbf{0}$$

    has non-trivial solution, where $c_1, c_2, \ldots, c_k$ are variables.

Linear independence (Cont.)

How to determine whether a subset is linearly independent or not?

- Let $S' \subset S \subset \mathbb{R}^n$,
  - if S' is linearly dependent, then S is linearly dependent;
  - if S is linearly independent, then S' is linearly independent;
- Let $S = \{\boldsymbol{u}\} \subset \mathbb{R}^n$, then S is linearly dependent iff $\boldsymbol{u} = \mathbf{0}$;
- Let $S = \{\boldsymbol{u}, \boldsymbol{v}\} \subset \mathbb{R}^n$, then S is linearly dependent iff $\boldsymbol{u} = a\boldsymbol{v}$ for some $a \in \mathbb{R}$ or $\boldsymbol{v} = b\boldsymbol{u}$ for some $b \in \mathbb{R}$;
- Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\} \subset \mathbb{R}^n$ where $k \geq 2$. Then
  - S is linearly dependent iff at least one vector $\boldsymbol{u}_i \in S$ can be written as a linear combination of other vectors in S;
  - S is linearly independent iff no vector in S can be written as a linear combination of other vectors in S.
- Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\} \subset \mathbb{R}^n$. If $k > n$, then S is linearly dependent.
- In $\mathbb{R}^n$, 2 vectors $\boldsymbol{u}, \boldsymbol{v}$ are linearly dependent iff they lie on the same line.
- In $\mathbb{R}^n$, 3 vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are linearly dependent iff they lie on the same plane.

Exercise (3.6)

Determine which of the following are subspaces of $\mathbb{R}^3$. Justify your answer.

(a) $\{(0,0,0)\}$.

(b) $\{(1,1,1)\}$.

(c) $\{(0,0,0),(1,1,1)\}$.

(d) $\{(0,0,c) \mid c$ *is an integer*$\}$.

(e) $\{(0,0,c) \mid c$ *is a non-negative real number*$\}$.

(f) $\{(0,0,c) \mid c$ *is a real number*$\}$.

(g) $\{(1,1,c) \mid c$ *is a real number*$\}$.

(h) $\{(a,b,c) \mid a, b, c$ *are real numbers and* $abc = 0\}$.

(i) $\{(a,b,c) \mid a, b, c$ *are real numbers and* $a \geq b \geq c\}$.

(j) $\{(a,b,c) \mid a, b$ *are real numbers and* $4a = 3b\}$.

(k) $\{(a,b,b) \mid a, b$ *are real numbers*$\}$.

(l) $\{(a,b,ab) \mid a, b$ *are real numbers*$\}$.

(m) $\{(a^2,b^2,c^2) \mid a, b, c$ *are real numbers*$\}$.

(n) $\{(a^3,b^3,c^3) \mid a, b, c$ *are real numbers*$\}$.

Recall

Methods for proving or disproving subspace

Solution.

(a) Yes. $\{\mathbf{0}\} = \mathrm{span}\{\mathbf{0}\}$ is a subspace of $\mathbb{R}^n$.

(b) No. It does not contain the zero vector.

(c) No. $(1,1,1)$ belongs to the set but $2(1,1,1)$ does not.

(d) No. $(0,0,1)$ belongs to the set but $\frac{1}{2}(0,0,1)$ does not.

(e) No. $(0,0,1)$ belongs to the set but $-(0,0,1)$ does not.

(f) Yes. It is $\mathrm{span}\{(0,0,1)\}$.

(g) No. It does not contain the zero vector.

(h) No. $(1,1,0)$ and $(0,0,1)$ belong to the set but $(1,1,0)+(0,0,1)=(1,1,1)$ does not.

(i) No. $(3,2,1)$ belongs to the set but $-(3,2,1)$ does not.

(j) Yes. It is a solution set of a homogeneous linear system.

(k) Yes. It is $\mathrm{span}\{(1,0,0),(0,1,1)\}$.

(l) No. $(1,1,1)$ belongs to the set but $2(1,1,1)$ does not.

(m) No. $(1,1,1)$ belongs to the set but $-(1,1,1)$ does not.

(n) Yes. It is $\mathbb{R}^3$, and hence a subspace.

$\square$

Exercise (3.12)

Let $\boldsymbol{A}$ be an $n \times n$ matrix. Define $V = \{\boldsymbol{u} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{u} = \boldsymbol{u}\}$.

(a) *Show that V is a subspace of $\mathbb{R}^n$.*

(b) *Let $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Write down the subspace V explicitly.*

Proof and Solution.

(a) Since $\boldsymbol{A}\boldsymbol{u} = \boldsymbol{u} \Leftrightarrow (\boldsymbol{A} - \boldsymbol{I})\boldsymbol{u} = \boldsymbol{0}$, V is the solution set of the homogeneous system $(\boldsymbol{A} - \boldsymbol{I})\boldsymbol{u} = 0$. By Theorem 3.2.9, V is a subspace of $\mathbb{R}^n$.

(b) $\boldsymbol{A} - \boldsymbol{I} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}$. A general solution of $\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ is $x = s$, $y = t$, $z = 0$, where $s, t \in \mathbb{R}$. So $V = \{(s, t, 0) \mid s, t \in \mathbb{R}\}$, i.e. V is the xy-plane in $\mathbb{R}^3$.

$\square$

Exercise (3.16)

Let $\boldsymbol{u}_1 = (2, 0, 2, -4)$, $\boldsymbol{u}_2 = (1, 0, 2, 5)$, $\boldsymbol{u}_3 = (0, 3, 6, 9)$, $\boldsymbol{u}_4 = (1, 1, 2, -1)$, $\boldsymbol{v}_1 = (-1, 2, 1, 0)$, $\boldsymbol{v}_2 = (3, 1, 4, 0)$, $\boldsymbol{v}_3 = (0, 1, 1, 3)$, $\boldsymbol{v}_4 = (-4, 3, -1, 6)$. *Determine if the following are true.*

(a) $\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\} \subset \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$.

(b) $\operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\} \subset \operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\}$.

(c) $\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\} = \mathbb{R}^4$.

(d) $\operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\} = \mathbb{R}^4$.

Method

Apply the method in Example 3.2.12.

Solution.

(a) By Gaussian elimination $\left(\begin{array}{cccc|c|c|c|c} -1 & 3 & 0 & -4 & 2 & 1 & 0 & 1 \\ 2 & 1 & 1 & 3 & 0 & 0 & 3 & 1 \\ 1 & 4 & 1 & -1 & 2 & 2 & 6 & 2 \\ 0 & 0 & 3 & 6 & -4 & 5 & 9 & -1 \end{array}\right) \rightarrow$

$\left(\begin{array}{cccc|c|c|c|c} -1 & 3 & 0 & -4 & 2 & 1 & 0 & 1 \\ 0 & 7 & 1 & -5 & 4 & 2 & 3 & 3 \\ 0 & 0 & 3 & 6 & -4 & 5 & 9 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 & 0 \end{array}\right)$, we know

$\boldsymbol{u}_2, \boldsymbol{u}_3 \notin \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$, and hence

$\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\} \not\subset \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$

Solution (Cont.)

(b) By Gaussian elimination $\left(\begin{array}{cccc|c|c|c|c} 2 & 1 & 0 & 1 & -1 & 3 & 0 & -4 \\ 0 & 0 & 3 & 1 & 2 & 1 & 1 & 3 \\ 2 & 2 & 6 & 2 & 1 & 4 & 1 & -1 \\ 0 & 5 & 9 & -1 & 0 & 0 & 3 & 6 \end{array}\right) \rightarrow$
$\left(\begin{array}{cccc|c|c|c|c} 2 & 1 & 0 & 1 & -1 & 3 & 0 & -4 \\ 0 & 1 & 6 & 1 & 2 & 1 & 1 & 3 \\ 0 & 0 & 3 & 1 & 2 & 1 & 1 & 3 \\ 0 & 0 & 0 & 1 & 4 & 2 & 5 & 12 \end{array}\right)$, we know
$\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4 \in \operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\}$, and hence $\operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\} \subset \operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\}$.

(c) Based on the same process, for any vector in $\mathbb{R}^4$, it is a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4$, and hence $\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\} = \mathbb{R}^4$.

(d) Based on the same process, there exists a vector in $\mathbb{R}^4$, which is not a linear combination of $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4$, and hence $\operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\} \neq \mathbb{R}^4$.

□

Exercise (3.18)

Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ *be the vectors and let*

$$S_1 = \{\boldsymbol{u}, \boldsymbol{v}\},\ S_2 = \{\boldsymbol{u}-\boldsymbol{v}, \boldsymbol{v}-\boldsymbol{w}, \boldsymbol{w}-\boldsymbol{u}\},\ S_3 = \{\boldsymbol{u}-\boldsymbol{v}, \boldsymbol{v}-\boldsymbol{w}, \boldsymbol{u}+\boldsymbol{w}\},$$

$$S_4 = \{\boldsymbol{u}, \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}+\boldsymbol{w}\},\ S_5 = \{\boldsymbol{u}+\boldsymbol{v}, \boldsymbol{v}+\boldsymbol{w}, \boldsymbol{u}+\boldsymbol{w}, \boldsymbol{u}+\boldsymbol{v}+\boldsymbol{w}\}.$$

(a) *Suppose* $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ *are vectors in* $\mathbb{R}^3$ *such that* $\operatorname{span}\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\} = \mathbb{R}^3$*. Determine which of the sets above span* $\mathbb{R}^3$*.*

(b) *Suppose* $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ *are linearly independent vectors in* $\mathbb{R}^n$*. Determine which of the sets above are linearly independent.*

Solution of (a).

- Note that $\text{span}(S_1)$ is a plane in $\mathbb{R}^3$. So S_1 does not span $\mathbb{R}^3$.
- Since $\boldsymbol{w} - \boldsymbol{u} = -(\boldsymbol{u} - \boldsymbol{v}) - (\boldsymbol{v} - \boldsymbol{w})$, $\text{span}(S_2) = \text{span}\{\boldsymbol{u} - \boldsymbol{v}, \boldsymbol{v} - \boldsymbol{w}\}$ which is also a plane in $\mathbb{R}^3$. So S_2 does not span $\mathbb{R}^3$.
- Note that $\text{span}(S_3) \subset \mathbb{R}^3$, and

$$\begin{aligned}
\boldsymbol{u} &= \frac{1}{2}\big[(\boldsymbol{u} - \boldsymbol{v}) + (\boldsymbol{v} - \boldsymbol{w}) + (\boldsymbol{u} + \boldsymbol{w})\big], \\
\boldsymbol{v} &= \frac{1}{2}\big[-(\boldsymbol{u} - \boldsymbol{v}) + (\boldsymbol{v} - \boldsymbol{w}) + (\boldsymbol{u} + \boldsymbol{w})\big], \\
\boldsymbol{w} &= \frac{1}{2}\big[-(\boldsymbol{u} - \boldsymbol{v}) - (\boldsymbol{v} - \boldsymbol{w}) + (\boldsymbol{u} + \boldsymbol{w})\big].
\end{aligned}$$

  Thus $\mathbb{R}^3 = \text{span}\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\} \subset \text{span}\{\boldsymbol{u} - \boldsymbol{v}, \boldsymbol{v} - \boldsymbol{w}, \boldsymbol{u} + \boldsymbol{w}\}$, and hence $\text{span}(S_3) = \mathbb{R}^3$.
- Using the same argument as for S_3, we can show that both S_4 and S_5 also span $\mathbb{R}^3$.

$\square$

Solution of (b).

- If there exist $a, b \in \mathbb{R}$, which are not both 0, such that $a\boldsymbol{u} + b\boldsymbol{v} = \boldsymbol{0}$, then $a\boldsymbol{u} + b\boldsymbol{v} + 0\boldsymbol{w} = \boldsymbol{0}$, i.e. $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are linearly dependent, contradiction.
- Since $(\boldsymbol{u} - \boldsymbol{v}) + (\boldsymbol{v} - \boldsymbol{w}) + (\boldsymbol{w} - \boldsymbol{u}) = \boldsymbol{0}$, they are linearly dependent.
- Suppose $a(\boldsymbol{u} - \boldsymbol{v}) + b(\boldsymbol{v} - \boldsymbol{w}) + c(\boldsymbol{w} + \boldsymbol{u}) = \boldsymbol{0}$, it is equivalent to $(a + c)\boldsymbol{u} + (-a + b)\boldsymbol{v} + (-b + c)\boldsymbol{w} = \boldsymbol{0}$. Since $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are linearly independent, we have $\begin{cases} a + c = 0 \\ -a + b = 0 \\ -b + c = 0 \end{cases}$. It is obvious that this system has only trivial solution, i.e. $\boldsymbol{u} - \boldsymbol{v}, \boldsymbol{v} - \boldsymbol{w}, \boldsymbol{w} + \boldsymbol{u}$ are linearly independent.
- By similarly method, we have that S_4 is linearly independent.
- Since $(\boldsymbol{u} + \boldsymbol{v}) + (\boldsymbol{v} + \boldsymbol{w}) + (\boldsymbol{u} + \boldsymbol{w}) - 2(\boldsymbol{u} + \boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{0}$, S_5 is linearly independent.

□

Exercise (3.22)

Let V and W be subspaces of $\mathbb{R}^n$.

(a) *Show that $V \cap W$ is a subspace of $\mathbb{R}^n$.*

(b) *Give an example of V and W in $\mathbb{R}^2$ such that $V \cup W$ is not a subspace.*

(c) *Show that $V \cup W$ is a subspace of $\mathbb{R}^n$ iff $V \subset W$ or $W \subset V$.*

Proof of part (a) and Solution of part (b).

(a) We use the result of Exercise 3.21 to prove that $V \cap W$ is a subspace of $\mathbb{R}^n$:

(1) Since both V and W contain the zero vector, the zero vector is contained in $V \cap W$ and hence $V \cap W$ is nonempty.

(2) Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be any two vectors in $V \cap W$ and let a and b be any real numbers. Since $\boldsymbol{u}$ and $\boldsymbol{v}$ are contained in V, $a\boldsymbol{u} + b\boldsymbol{v}$ is also contained in V. Similarly, $a\boldsymbol{u} + b\boldsymbol{v}$ is also contained in W. Thus $a\boldsymbol{u} + b\boldsymbol{v}$ is contained in $V \cap W$.

By the result of Exercise 3.21, $V \cap W$ is a subspace of $\mathbb{R}^n$.

(b) Let $V = \{(x, 0) \mid x \in \mathbb{R}\}$ and $W = \{(0, y) \mid y \in \mathbb{R}\}$. Then both V and W are lines through the origin and hence are subspaces of $\mathbb{R}^n$. But $V \cap W$ is a union of two lines which is not a subspace of $\mathbb{R}^n$.

$\square$

Proof of part (c).

($\Leftarrow$) If $V \subset W$, then $V \cup W = W$ is a subspace of $\mathbb{R}^n$; if $W \subset V$, then $W \cup V = V$ is a subspace of $\mathbb{R}^n$.

($\Rightarrow$)
(1) Suppose $V \not\subset W$. We want to show that $W \subset V$.
(2) Take any vector $\boldsymbol{x} \in W$, we want to show $\boldsymbol{x} \in V$.
(3) Since $V \not\subset W$, there exists a vector $\boldsymbol{y} \in V$ but $\boldsymbol{y} \notin W$.
(4) Since $V \cup W$ is a subspace of $\mathbb{R}^n$ and $\boldsymbol{x}, \boldsymbol{y} \in V \cup W$, we have $\boldsymbol{x} + \boldsymbol{y} \in V \cup W$, i.e. either $\boldsymbol{x} + \boldsymbol{y} \in V$ or $\boldsymbol{x} + \boldsymbol{y} \in W$.
(5) If $\boldsymbol{x} + \boldsymbol{y} \in W$. As W is a subspace of $\mathbb{R}^n$, we have $\boldsymbol{y} = (\boldsymbol{x} + \boldsymbol{y}) - \boldsymbol{x} \in W$ which contradict that $\boldsymbol{y} \notin W$ as mentioned above.
(6) Now we know that $\boldsymbol{x} + \boldsymbol{y} \in V$. As V is a subspace of $\mathbb{R}^n$, we have $\boldsymbol{x} = (\boldsymbol{x} + \boldsymbol{y}) - \boldsymbol{y} \in V$.
(7) Since every vector in W must be contained in V, $W \subset V$.

$\square$

Exercise (3.23)

(All vectors in this question are column vectors.) Let $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ be vectors in $\mathbb{R}^n$ and $\boldsymbol{A}$ an $n \times n$ matrix.

(a) *Show that if $\boldsymbol{A}\boldsymbol{u}_1, \boldsymbol{A}\boldsymbol{u}_2, \ldots, \boldsymbol{A}\boldsymbol{u}_k$ are linearly independent, then $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are linearly independent.*

(b) *Suppose $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are linearly independent.*
- *Show that if $\boldsymbol{A}$ is invertible, then $\boldsymbol{A}\boldsymbol{u}_1, \boldsymbol{A}\boldsymbol{u}_2, \ldots, \boldsymbol{A}\boldsymbol{u}_k$ are linearly independent.*
- *If $\boldsymbol{A}$ is not invertible, are $\boldsymbol{A}\boldsymbol{u}_1, \boldsymbol{A}\boldsymbol{u}_2, \ldots, \boldsymbol{A}\boldsymbol{u}_k$ linearly independent?*

Proof of part (a).

Suppose $c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k = \boldsymbol{0}$, then

$$c_1\boldsymbol{A}\boldsymbol{u}_1 + c_2\boldsymbol{A}\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{A}\boldsymbol{u}_k = \boldsymbol{A}(c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k) = \boldsymbol{0}.$$

Since $\boldsymbol{A}\boldsymbol{u}_1, \boldsymbol{A}\boldsymbol{u}_2, \ldots, \boldsymbol{A}\boldsymbol{u}_k$ are linearly independent, we have $c_1 = c_2 = \cdots = c_k = 0$, i.e. $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are linearly independent. $\square$

Proof of part (b).

- (1) Suppose $c_1\boldsymbol{A}\boldsymbol{u}_1 + c_2\boldsymbol{A}\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{A}\boldsymbol{u}_k = 0$, then $\boldsymbol{A}(c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k) = \boldsymbol{0}$.
  (2) Since $\boldsymbol{A}$ is invertible, $c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k = \boldsymbol{0}$.
  (3) Since $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are linearly independent, we have $c_1 = c_2 = \cdots = c_k = 0$, and hence $\boldsymbol{A}\boldsymbol{u}_1, \boldsymbol{A}\boldsymbol{u}_2, \ldots, \boldsymbol{A}\boldsymbol{u}_k$ are linearly independent.
- No conclusion. For example, let $\boldsymbol{u}_1 = (1,0,0)^T$ and $\boldsymbol{u}_2 = (0,1,0)^T$: It is obvious that $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are linearly independent.
  - If $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, then $\boldsymbol{A}\boldsymbol{u}_1$ and $\boldsymbol{A}\boldsymbol{u}_2$ are linearly independent.
  - If $\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, then $\boldsymbol{A}\boldsymbol{u}_1$ and $\boldsymbol{A}\boldsymbol{u}_2$ are linearly dependent.

□

Exercise (3.24)

Determine which of the following statements are true. Justify your answer.

(a) $\mathbb{R}^2$ *is a subspace of* $\mathbb{R}^3$.

(b) *The solution set of* $x + 2y - z = 0$ *is a subspace of* $\mathbb{R}^3$.

(c) *The solution set of* $x + 2y - z = 1$ *is a subspace of* $\mathbb{R}^3$.

(d) *If* $\boldsymbol{u}, \boldsymbol{v}$ *are nonzero vectors in* $\mathbb{R}^2$ *such that* $\boldsymbol{u} \neq \boldsymbol{v}$, *then* $\mathrm{span}\{\boldsymbol{u}, \boldsymbol{v}\} = \mathbb{R}^2$.

(e) *If* S_1 *and* S_2 *are two subsets of a vector space, then* $\mathrm{span}(S_1 \cap S_2) = \mathrm{span}(S_1) \cap \mathrm{span}(S_2)$.

(f) *If* S_1 *and* S_2 *are two subsets of a vector space, then* $\mathrm{span}(S_1 \cup S_2) = \mathrm{span}(S_1) \cup \mathrm{span}(S_2)$.

(g) *If* S_1 *and* S_2 *are two subsets of a vector space, then* $\mathrm{span}(S_1 \cup S_2) = \mathrm{span}(S_1) + \mathrm{span}(S_2)$.

Proof.

(a) False. $\mathbb{R}^2$ is not even a subset of $\mathbb{R}^3$. (We can only say that the xy-plane $\{(x, y, 0) \mid x, y \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^3$.)

(b) The equation $x + 2y - z = 0$ forms a homogeneous system of linear equations.

(c) False. Note that $(0, 0, 0)$ is not a solution of $ax + by + cz = 1$.

(d) False. For example, let $\boldsymbol{u} = (1, 1), \boldsymbol{v} = (2, 2)$.

(e) False. For example, let $S_1 = \{(1, 0), (0, 1)\}, S_2 = \{(1, 0), (0, 2)\}$.

(f) False. For example, let $S_1 = \{(1, 0)\}, S_2 = \{(0, 1)\}$.

(g) True.

- For any element $\boldsymbol{u}$ of $\mathrm{span}(S_1 \cup S_2)$, it can be expressed as a linear combination of $S_1 \cup S_2$. Hence, $\boldsymbol{u} = \boldsymbol{u}_1 + \boldsymbol{u}_2$ where $\boldsymbol{u}_1 \in \mathrm{span}(S_1)$ and $\boldsymbol{u}_2 \in \mathrm{span}(S_2)$.
- For any elements $\boldsymbol{u}_1 \in \mathrm{span}(S_1)$ and $\boldsymbol{u}_2 \in \mathrm{span}(S_2)$, $\boldsymbol{u}_1 + \boldsymbol{u}_2$ is a linear combination of $S_1 \cup S_2$.

$\square$

Exercise (3.10(a))

Let V and W be subspaces of $\mathbb{R}^n$. Define $V + W = \{\boldsymbol{v} + \boldsymbol{w} \mid \boldsymbol{v} \in V, \boldsymbol{w} \in W\}$. Then $V + W$ is a subspace of $\mathbb{R}^n$.

Proof.

We use the result of Exercise 3.21 to prove that $V + W$ is a subspace of $\mathbb{R}^n$:

(1) Since both V and W contain the zero vector, the zero vector is contained in $V + W$ and hence $V + W$ is nonempty.

(2) Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be any two vectors in $V + W$ and let a and b be any real numbers. Then $\boldsymbol{u}$ and $\boldsymbol{v}$ can be expressed as $\boldsymbol{u} = \boldsymbol{u}_1 + \boldsymbol{u}_2$ and $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$, where $\boldsymbol{u}_1, \boldsymbol{v}_1 \in V$ and $\boldsymbol{u}_2, \boldsymbol{v}_2 \in W$.

$$a\boldsymbol{u} + b\boldsymbol{v} = (a\boldsymbol{u}_1 + b\boldsymbol{v}_1) + (a\boldsymbol{u}_2 + b\boldsymbol{v}_2)$$

is contained in $V + W$.

By the result of Exercise 3.21, $V + W$ is a subspace of $\mathbb{R}^n$. □

Exercise (3.20)

Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ be vectors in $\mathbb{R}^3$ such that $V = \operatorname{span}\{\boldsymbol{u}, \boldsymbol{v}\}$ and $W = \operatorname{span}\{\boldsymbol{u}, \boldsymbol{w}\}$ are planes in $\mathbb{R}^3$. Find $V \cap W$ if

(a) *$\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are linearly independent.*

(b) *$\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are not linearly independent.*

Solution.

(a) If $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ are linearly independent, then the two planes V and W intersect at the line spanned by $\boldsymbol{u}$ and hence $V \cap W = \operatorname{span}\{\boldsymbol{u}\}$.

(b) V and W are planes in $\mathbb{R}^3$. So $\boldsymbol{u}, \boldsymbol{v}$ are linearly independent and $\boldsymbol{u}, \boldsymbol{w}$ are linearly independent. If $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are linearly dependent, then $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ must lie on the same plane and hence $V = W = V \cap W$.

$\square$

Exercise (3.21)

Let V be a non-empty subset of $\mathbb{R}^n$. Show that V is a subspace iff for any $\boldsymbol{u}, \boldsymbol{v} \in V$ and $a, b \in \mathbb{R}$, $a\boldsymbol{u} + b\boldsymbol{v} \in V$.

Proof.

($\Rightarrow$) If V is a subspace of $\mathbb{R}^n$, then by Theorem 3.2.5.2, for any $\boldsymbol{u}, \boldsymbol{v} \in V$ and $a, b \in \mathbb{R}$, $a\boldsymbol{u} + b\boldsymbol{v} \in V$.

($\Leftarrow$)
1. Suppose for any $\boldsymbol{u}, \boldsymbol{v} \in V$ and $a, b \in \mathbb{R}$, $a\boldsymbol{u} + b\boldsymbol{v} \in V$.
2. Take $a = b = 0$, then we know that zero vector $\mathbf{0} \in V$.
3. If $V = \{\mathbf{0}\}$, then V is a subspace of $\mathbb{R}^n$, see Remark 3.2.4.1.
4. Suppose $V \neq \{\mathbf{0}\}$. Since V is a non-empty subset of $\mathbb{R}^n$, it has at least 1 and at most n linearly independent vectors, see Theorem 3.3.9.
5. Let S be a largest set of linearly independent vectors in V. Then $\mathrm{span}(S) = V$; if not, there exists $\boldsymbol{v} \in V$ but $\boldsymbol{v} \notin \mathrm{span}(S)$, and by Problem 3.3.11, $S \cup \{\boldsymbol{v}\}$ is linearly independent which violates our assumption on S
6. So V is a subspace of $\mathbb{R}^n$.

$\square$

Change log

- Page 109: Add a slide for "operations of subspaces";
- Page 114: Revise a typo: "$m \times n$" to "$n \times n$";
- Page 119: Revise a mistake.

Last modified: 11:45, March 4, 2011.

Schedule of Tutorial 6

- Any question about last tutorial
- Review concepts: Vector spaces:
  - Bases, coordinate vector relative to a basis;
  - Dimension;
  - Transition matrix.
- Tutorial: 3.25, 3.28, 3.31, 3.32, 3.33, 3.36
- Additional material:
  - 3.35, 3.37;
  - Question 6 in Final 2001–2002(I);
  - Question 3 in Final 2001–2002(II);
  - Question 4 in Final 2004–2005(II);
  - Question 6(C) in Final 2005–2006(I);
  - Question 1(C) in Final 2008–2009(II).

Bases

- Let $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ be linearly independent vectors in $\mathbb{R}^n$. Suppose $\boldsymbol{u}_{k+1}$ is a vector in $\mathbb{R}^n$, and not a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$. Then $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k, \boldsymbol{u}_{k+1}$ are linearly independent.
- Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ be a subset of a vector space V. S is called a basis for V if
  1. S is linearly independent;
  2. S spans V.
- A basis for a vector space V contains the smallest possible number of vectors that can span V.
- Existence of bases: Problem 3.4.8:
  1. Suppose $S \subset V$ and $\mathrm{span}(S) = V$, then there exists $S' \subset S$, such that S' is a basis for V. (Remove "redundant" vectors form S repeatedly.)
  2. Suppose T is a set of linearly independent vectors in V. Then there exists a basis T' for V such that $T \subset T'$. (Add in suitable vectors to T repeatedly.)

Bases (Cont.)

- Theorem 3.4.5: Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ be a basis for a vector space V, then every vector $\boldsymbol{v} \in V$ can be expressed in the form $\boldsymbol{v} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k$ in exactly one way, where $c_1, c_2, \ldots, c_k \in \mathbb{R}$.
- Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ be a basis for a vector space V and $\boldsymbol{v} \in V$.
  - If $\boldsymbol{v} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k$, then the coefficients $c_1, c_2, \ldots, c_k$ are called the coordinates of $\boldsymbol{v}$ relative to the basis S.
  - The vector $(\boldsymbol{v})_S = (c_1, c_2, \ldots, c_k) \in \mathbb{R}^k$ is called the coordinate vector of $\boldsymbol{v}$ relative to the basis S.
- Except the zero space, any vector space has infinitely many different bases. For example, for any $x \neq 0$, $\{(x, 0), (0, x)\}$ is a basis for $\mathbb{R}^2$.
- $S = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$, where $\boldsymbol{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^n$, is the standard basis for $\mathbb{R}^n$, and we have

$$(\boldsymbol{u})_S = (u_1, u_2, \ldots, u_n) = \boldsymbol{u}.$$

Dimension

- Theorem 3.5.1: Let V be a vector space which has a basis with k vectors. Then
  - any subset of V with more than k vectors is always linearly dependent;
  - any subset of V with less than k vectors can not span V.
- Definition 3.5.3: The dimension of a vector space V, denoted by $\dim(V)$, is defined to be the number of vectors in a basis for V. In addition, we define the dimension of the zero space to be zero.
- $\dim(\mathbb{R}^n) = n$ for any $n \in \mathbb{N}$.
- Theorem 3.5.6: Let V be a vector space of dimension k and S a subset of V. The following are equivalent:
  (1) S is a basis for V;
  (2) S is linearly independent, and $|S| = k = \dim(V)$;
  (3) S spans V, and $|S| = k = \dim(V)$.
- Theorem 3.5.8: Let $\boldsymbol{A}$ be an $n \times n$ matrix. The following statements are equivalent:
  (1) $\boldsymbol{A}$ is invertible;
  (2) The linear system $\boldsymbol{A}\boldsymbol{x} = 0$ has only trivial solution;
  (3) The RREF of $\boldsymbol{A}$ is an identity matrix;
  (4) $\boldsymbol{A}$ can be expressed as a product of elementary matrices;
  (5) $\det(\boldsymbol{A}) \neq 0$;
  (6) The rows of $\boldsymbol{A}$ form a basis for $\mathbb{R}^n$;
  (7) The columns of $\boldsymbol{A}$ form a basis for $\mathbb{R}^n$.

How to

- How to prove S to be a basis for a vector space V:

  (1) $S \subset V$;
  (2-1) S is linearly independent;
  (2-2) S spans V;
  (2-3) $|S| = \dim(V)$.

  If we show that the Condition (1) and any two of the Conditions (2-1), (2-2) and (2-3) are satisfied, then S is a basis for V.
- For $\mathbb{R}^n$, if $|\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n| \neq 0$, then $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ is a basis for $\mathbb{R}^n$.
- How to find a basis for a subspace V: express a general vector in V as a linear combination.
- How to compute dimension for a vector space: find a basis first.

Transition matrices

- Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_k\}$ be a basis for a vector space V and $\boldsymbol{v}$ be a vector in V. If $\boldsymbol{v} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k$, then the vectors

$$(\boldsymbol{v})_S = (c_1, c_2, \dots, c_k), \quad [\boldsymbol{v}]_S = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$$

  are called the coordinate vector of $\boldsymbol{v}$ relative to S.

- Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_k\}$ and $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_k\}$ be two bases for a vector space V. Then for any $\boldsymbol{w} \in V$, we have

$$[\boldsymbol{w}]_T = \left([\boldsymbol{u}_1]_T, [\boldsymbol{u}_2]_T, \dots, [\boldsymbol{u}_k]_T\right)[\boldsymbol{w}]_S.$$

  Hence the matrix

$$\boldsymbol{P} = \left([\boldsymbol{u}_1]_T, [\boldsymbol{u}_2]_T, \dots, [\boldsymbol{u}_k]_T\right)$$

  is called the transition matrix from S to T.

- Let S and T be two bases of a vector space and let $\boldsymbol{P}$ be the transition matrix from S to T. Then
  - $\boldsymbol{P}$ is invertible;
  - $\boldsymbol{P}^{-1}$ is the transition matrix from T to S.

Exercise (3.25)

Determine which of the following sets are bases for $\mathbb{R}^3$.

(a) $S_1 = \{(1, 0, -1), (-1, 2, 3)\}$.

(b) $S_2 = \{(1, 0, -1), (-1, 2, 3), (0, 3, 0)\}$.

(c) $S_3 = \{(1, 0, -1), (-1, 2, 3), (0, 3, 3)\}$.

(d) $S_4 = \{(1, 0, -1), (-1, 2, 3), (0, 3, 0), (1, -1, 1)\}$.

Solution.

(a) No. There are too few vectors. $(|S_1| = 2 < 3 = \dim(\mathbb{R}^3))$

(b) Yes. $S_2 \subset \mathbb{R}^3$, S_2 is linearly independent (easy to check), and $|S_2| = 3 = \dim(\mathbb{R}^3)$.

(c) No. S_3 is linearly dependent: $3(1, 0, -1) + 3(-1, 2, 3) - 2(0, 3, 3) = (0, 0, 0)$.

(d) No. There are too many vectors. $(|S_4| = 4 > 3 = \dim(\mathbb{R}^3))$

$\square$

Exercise (3.28)

Let $V = \{(a+b, a+c, c+d, b+d) \mid a, b, c, d \in \mathbb{R}\}$ *and* $S = \{(1,1,0,0), (1,0,-1,0), (0,-1,0,1)\}$.

(a) *Show that* V *is a subspace of* $\mathbb{R}^4$ *and* S *is a basis for* V.

(b) *Find the coordinate vector of* $\boldsymbol{u} = (1,2,3,2)$ *relative to* S.

(c) *Find a vector* $\boldsymbol{v}$ *such that* $(\boldsymbol{v})_S = (1,3,-1)$.

Proof.

(a) $V = \{a(1,1,0,0) + b(1,0,0,1) + c(0,1,1,0) + d(0,0,1,1) \mid a, b, c, d \in \mathbb{R}\} = \text{span}\{(1,1,0,0), (1,0,0,1), (0,1,1,0), (0,0,1,1)\}$ and hence is a subspace of $\mathbb{R}^4$. It is easy to see that $S \subset V$, S is linearly independent and

$$\text{span}(S) = \text{span}\{(1,1,0,0), (1,0,0,1), (0,1,1,0), (0,0,1,1)\} = V.$$

So S is a basis for V.

(b) Let $(1,2,3,2) = c_1(1,1,0,0) + c_2(1,0,-1,0) + c_3(0,-1,0,1)$. Then we will get $c_1 = 4$, $c_2 = -3$ and $c_3 = 2$, that is, the coordinate vector of $\boldsymbol{u}$ relative to S is $(4,-3,2)$.

(c) $\boldsymbol{v} = 1(1,1,0,0) + 3(1,0,-1,0) - 1(0,-1,0,1) = (4,2,-3,-1)$.

□

Exercise (3.31)

Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ be a basis for a vector space V. Determine whether $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is a basis for V if

(a) $\boldsymbol{v}_1 = \boldsymbol{u}_1,\ \boldsymbol{v}_2 = \boldsymbol{u}_1 + \boldsymbol{u}_2,\ \boldsymbol{v}_3 = \boldsymbol{u}_1 + \boldsymbol{u}_2 + \boldsymbol{u}_3$.

(b) $\boldsymbol{v}_1 = \boldsymbol{u}_1 - \boldsymbol{u}_2,\ \boldsymbol{v}_2 = \boldsymbol{u}_2 - \boldsymbol{u}_3,\ \boldsymbol{v}_3 = \boldsymbol{u}_3 - \boldsymbol{u}_1$.

Solution.

(a)
- $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3 \in V$, since V is a vector space;
- (1) Suppose $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3 = \boldsymbol{0}$. Then

  $$(c_1 + c_2 + c_3)\boldsymbol{u}_1 + (c_2 + c_3)\boldsymbol{u}_2 + c_3\boldsymbol{u}_3 = \boldsymbol{0}.$$

  (2) Since $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ are linearly independent, $c_1 + c_2 + c_3 = c_2 + c_3 = c_3 = 0$, and hence $c_1 = c_2 = c_3 = 0$.

  (3) Hence, $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ are linearly independent.
- $\dim(V) = 3 = |\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}|$.

Hence, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is a basis for V.

(b) Since $\boldsymbol{v}_1 + \boldsymbol{v}_2 + \boldsymbol{v}_3 = 0$, they are linearly dependent. Hence, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is not a basis for V.

$\square$

Exercise (3.32)

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$*, where* $\boldsymbol{u}_1 = (3, -2, 5)$*,* $\boldsymbol{u}_2 = (1, -4, 4)$*,* $\boldsymbol{u}_3 = (0, 3, -2)$*.*

(a) *Show that* S *is a basis for* $\mathbb{R}^3$*.*

(b) *Show that* $T = \{\boldsymbol{u}_1 - \boldsymbol{u}_2, \boldsymbol{u}_1 + 2\boldsymbol{u}_2 - \boldsymbol{u}_3, \boldsymbol{u}_2 + 2\boldsymbol{u}_3\}$ *is also a basis for* $\mathbb{R}^3$*.*

(c) *Find the coordinate vector of* $\boldsymbol{v} = (1, 0, 1)$ *relative to* S*.*

(d) *Find a vector* $\boldsymbol{w}$ *in* $\mathbb{R}^3$ *such that* $(\boldsymbol{w})_T = (1, 0, 1)$*.*

(e) *Find the transition matrix from* T *to* S *and the transition matrix from* S *to* T*.*

(f) *Let* $\boldsymbol{x}$ *be a vector in* $\mathbb{R}^3$ *such that* $(\boldsymbol{x})_T = (1, 1, 2)$*. Find* $(\boldsymbol{x})_S$*.*

Proof of parts (a,b).

(a) Since $\begin{vmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{vmatrix} = -1$, by Theorem 3.5.8, S is a basis for $\mathbb{R}^3$.

(b)
1. $T \subset \mathbb{R}^3$;
2. Let $c_1(\boldsymbol{u}_1 - \boldsymbol{u}_2) + c_2(\boldsymbol{u}_1 + 2\boldsymbol{u}_2 - \boldsymbol{u}_3) + c_3(\boldsymbol{u}_2 + 2\boldsymbol{u}_3) = \boldsymbol{0}$. Then $(c_1 + c_2)\boldsymbol{u}_1 + (-c_1 + 2c_2 + c_3)\boldsymbol{u}_2 + (-c_2 + 2c_3)\boldsymbol{u}_3 = \boldsymbol{0}$. By part (a), S is linearly independent, thus $\begin{cases} c_1 + c_2 = 0 \\ -c_1 + 2c_2 + c_3 = 0 \\ -c_2 + 2c_3 = 0 \end{cases}$.
   The system has only the trivial solution. So T is linearly independent.
3. $\dim(\mathbb{R}^3) = 3 = |T|$.

Hence, T is a basis for $\mathbb{R}^3$.

Solution of parts (c–f).

(c) Let $\boldsymbol{v} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + c_3\boldsymbol{u}_3$. Then we need to solve the linear system:

$$\begin{cases} 3c_1 + c_2 = 1 \\ -2c_1 - 4c_2 + 3c_3 = 0. \\ 5c_1 + 4c_2 - 2c_3 = 1 \end{cases}$$

By Gaussian elimination, we get $c_1 = 1$, $c_2 = -2$ and $c_3 = -2$. Hence $(\boldsymbol{v})_S = (1, -2, -2)$.

(d) $\boldsymbol{w} = 1(\boldsymbol{u}_1 - \boldsymbol{u}_2) + 0(\boldsymbol{u}_1 + 2\boldsymbol{u}_2 - \boldsymbol{u}_3) + 1(\boldsymbol{u}_2 + 2\boldsymbol{u}_3) = (3, 4, 1)$.

(e) The transition matrix from T to S is $\boldsymbol{P} = \big([\boldsymbol{v}_1]_S, [\boldsymbol{v}_2]_S, [\boldsymbol{v}_3]_S\big)$. Since $[\boldsymbol{v}_1]_S = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, $[\boldsymbol{v}_2]_S = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ and $[\boldsymbol{v}_3]_S = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$. Thus $\boldsymbol{P} = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix}$, and hence the transition matrix from S to T is $\boldsymbol{P}^{-1} = \frac{1}{7}\begin{pmatrix} 5 & -2 & 1 \\ 2 & 2 & -1 \\ 1 & 1 & 3 \end{pmatrix}$. s

(f) $[\boldsymbol{x}]_S = \boldsymbol{P}[\boldsymbol{x}]_T = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}$, hence $(\boldsymbol{x})_S = [\boldsymbol{x}]_S^T = (2, 3, 3)$.

□

Exercise (3.33)

Let $V = \{(x, y, z) \mid 2x - y + z = 0\}$, $S = \{(0, 1, 1), (1, 2, 0)\}$, $T = \{(1, 1, -1), (1, 0, -2)\}$.

(a) *Show that both* S *and* T *are basis for* V.

(b) *Find the transition matrix from* T *to* S *and the transition matrix from* S *to* T.

(c) *Show that* $S' = S \cup \{(2, -1, 1)\}$ *is a basis for* $\mathbb{R}^3$.

Proof of parts (a,c).

(a)
- $S \subset V$.
- Since $V = \{(x, y, z) \mid 2x - y + z = 0\} = \{(x, 2x + z, z) \mid x, z \in \mathbb{R}\} = \mathrm{span}\{(1, 2, 0), (0, 1, 1)\} = \mathrm{span}(S)$, S spans V.
- It is obvious that S is linearly independent.

Hence, S is a basis for V.
Similarly, we have that T is linearly independent. Since $T \subset V$ and $\dim(V) = |S| = 2 = |T|$, T is also a basis for V.

(c) Since $(2, -1, 1)$ does not satisfy the equation $2x - y + z = 0$, it can not be expressed as a linear combination of S, i.e., S' is linearly independent. As $\dim(\mathbb{R}^3) = 3$, S' is a basis for $\mathbb{R}^3$.

□

Solution of part (b).

1. By Gauss-Jordan elimination, we have

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

2. Thus $[(1,1,-1)]_S = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $[(1,0,-2)]_S = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$.
3. The transition matrix from T to S is $\begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}$.
4. And hence the transition matrix from S to T is $\begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}$.

$\square$

Exercise (3.36)

Let V and W be subspaces of $\mathbb{R}^n$. Show that $\dim(V+W) = \dim(V) + \dim(W) - \dim(V \cap W)$.

Proof.

- Let $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ be a basis for $V \cap W$. By Problem 3.4.8.2, there exists vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m \in V$ such that $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ is a basis for V and there exists vectors $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n \in W$ such that $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ is a basis for W. It is easy to see that $V + W = \text{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$.
- Consider $a_1\boldsymbol{u}_1 + \cdots + a_k\boldsymbol{u}_k + b_1\boldsymbol{v}_1 + \cdots + b_m\boldsymbol{v}_m + c_1\boldsymbol{w}_1 + \cdots + c_n\boldsymbol{w}_n = \boldsymbol{0}$ $(*)$. Since $c_1\boldsymbol{w}_1 + \cdots + c_n\boldsymbol{w}_n = -(a_1\boldsymbol{u}_1 + \cdots + a_k\boldsymbol{u}_k + b_1\boldsymbol{v}_1 + \cdots + b_m\boldsymbol{v}_m) \in V \cap W$, there exists $d_1, \ldots, d_k \in \mathbb{R}$ such that $c_1\boldsymbol{w}_1 + \cdots + c_n\boldsymbol{w}_n = d_1\boldsymbol{u}_1 + \cdots + d_k\boldsymbol{u}_k$, i.e., $c_1\boldsymbol{w}_1 + \cdots + c_n\boldsymbol{w}_n - d_1\boldsymbol{u}_1 - \cdots - d_k\boldsymbol{u}_k = \boldsymbol{0}$. As $\{u_1, \ldots, u_k, w_1, \ldots, w_n\}$ is linearly independent, $c_1 = \cdots = c_n = d_1 = \cdots = d_k = 0$.
- Substituting $c_1 = \cdots = c_n = 0$ into $(*)$, we have $a_1\boldsymbol{u}_1 + \cdots + a_k\boldsymbol{u}_k + b_1\boldsymbol{v}_1 + \cdots + b_m\boldsymbol{v}_m = \boldsymbol{0}$. As $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ is linearly independent, $a_1 = \cdots = a_k = b_1 = \cdots = b_m = 0$.
- So $(*)$ has only the trivial solution and hence $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ is linearly independent. We have shown that $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ is a basis for $V + W$.
- Thus $\dim(V + W) = k + m + n = \dim(V) + \dim(W) - \dim(V \cap W)$.

$\square$

Exercise (3.35)

Let V be a vector space of dimension of n. Show that there exists $n+1$ vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n, \boldsymbol{u}_{n+1}$ such that every vector in V can be expressed as a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{n+1}$ with non-negative coefficients.

Proof.

- Take a basis $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ for V. Define $\boldsymbol{u}_{n+1} = -\boldsymbol{u}_1 - \boldsymbol{u}_2 - \cdots - \boldsymbol{u}_n$.
- For any $\boldsymbol{v} \in V$, $\boldsymbol{v} = a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2 + \cdots + a_n\boldsymbol{u}_n$ for some $a_1, a_2, \ldots, a_n \in \mathbb{R}$.
- Let $a = \min\{0, a_1, a_2, \ldots, a_n\}$. Then $\boldsymbol{v} = (a_1 - a)\boldsymbol{u}_1 + (a_2 - a)\boldsymbol{u}_2 + \cdots + (a_n - a)\boldsymbol{u}_n + (-a)\boldsymbol{u}_{n+1}$ where $a_i - a \geq 0$, for $i = 1, 2, \ldots, n$, and $-a \geq 0$.
- So every vector in V can be expressed as a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n, \boldsymbol{u}_{n+1}$ with non-negative coefficients.

$\square$

Exercise (3.37)

Determine which of the following statements are true. Justify your answer.

(a) *If S_1 and S_2 are basis for V and W respectively, where V and W are subspaces of a vector space, then $S_1 \cap S_2$ is a basis for $V \cap W$.*

(b) *If S_1 and S_2 are basis for V and W respectively, where V and W are subspaces of a vector space, then $S_1 \cup S_2$ is a basis for $V + W$.*

(c) *If V and W are subspace of a vector space, then there exists a basis S_1 for V and a basis S_2 for W such that $S_1 \cap S_2$ is a basis for $V \cap W$.*

(d) *If V and W are subspace of a vector space, then there exists a basis S_1 for V and a basis S_2 for W such that $S_1 \cup S_2$ is a basis for $V + W$.*

Solution.

(a) False. For example, let $S_1 = \{(1,0),(0,1)\}$ and $S_2 = \{(1,0),(0,2)\}$ where $V = W = \mathbb{R}^2$.

(b) False. For example, let $S_1 = \{(1,0)\}$ and $S_2 = \{(1,1),(0,1)\}$ where $V = \mathrm{span}(S_1)$ and $W = V + W = \mathbb{R}^2$. Note that $S_1 \cup S_2$ is linearly dependent.

(c) True. See the proof of Exercise 3.36.

(d) True. See the proof of Exercise 3.36.

□

Exercise (Question 6 in Final 2001–2002(I))

Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ *be a basis of* $\mathbb{R}^3$ *and let* $\boldsymbol{u}_1 = a\boldsymbol{v}_1 + b\boldsymbol{v}_2 + c\boldsymbol{v}_3$*,* $\boldsymbol{u}_2 = d\boldsymbol{v}_1 + e\boldsymbol{v}_2 + f\boldsymbol{v}_3$*,* $\boldsymbol{u}_3 = g\boldsymbol{v}_1 + h\boldsymbol{v}_2 + k\boldsymbol{v}_3$*. Suppose that*

$$\begin{pmatrix} a & d & g \\ b & e & h \\ c & f & k \end{pmatrix}$$

is invertible. Prove that $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ *is a basis of* $\mathbb{R}^3$*.*

Proof.

It is easy to see $\begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{pmatrix} \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & k \end{pmatrix}$.

Thus $\det \begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 \end{pmatrix} = \det \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{pmatrix} \det \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & k \end{pmatrix} \neq 0.$

Therefore $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ are linearly independent.
Since $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\} \subset \mathbb{R}^3$ and $|\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}| = 3 = \dim(\mathbb{R}^3)$, we have that $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is a basis of $\mathbb{R}^3$. □

Exercise (Question 3 in Final 2001–2002(II))

Let W be the real vector space of all 3×3 symmetric matrices. Find a basis of W. Justify your answers.

Solution.

1. $\boldsymbol{A} = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$ is the typical element of W.
2. $\boldsymbol{A} = a \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} +$
   $e \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + f \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
3. These 6 matrices are linearly independent (by definition).
4. Hence, the set consisting of these 6 matrices is a basis of W.

$\square$

Exercise (Question 4 in Final 2004—2005(II))

Let $\boldsymbol{A}$ *be a basis of* $\mathbb{R}^n$ *with* $\det(\boldsymbol{A}) = 0$ *and let* $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ *be a basis of* $\mathbb{R}^n$*. Prove that* $\{\boldsymbol{A}\boldsymbol{v}_1, \boldsymbol{A}\boldsymbol{v}_2, \ldots, \boldsymbol{A}\boldsymbol{v}_n\}$ *is linearly dependent.*

Proof.

Since

$$\begin{pmatrix} \boldsymbol{A}\boldsymbol{v}_1 & \boldsymbol{A}\boldsymbol{v}_2 & \cdots \boldsymbol{A}\boldsymbol{v}_n \end{pmatrix} = \boldsymbol{A} \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{pmatrix},$$

we have

$$\det \begin{pmatrix} \boldsymbol{A}\boldsymbol{v}_1 & \boldsymbol{A}\boldsymbol{v}_2 & \cdots \boldsymbol{A}\boldsymbol{v}_n \end{pmatrix} = \det(\boldsymbol{A}) \det \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{pmatrix} = 0.$$

Therefore $\{\boldsymbol{A}\boldsymbol{v}_1, \boldsymbol{A}\boldsymbol{v}_2, \ldots, \boldsymbol{A}\boldsymbol{v}_n\}$ is linearly dependent. □

Exercise (Question 6(c) in Final 2005–2006(I))

Let $S = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ be a basis for a vector space V. Show that

$$T = \{\boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{x}_2 + \boldsymbol{x}_3, \ldots, \boldsymbol{x}_{n-1} + \boldsymbol{x}_n, \boldsymbol{x}_n + \boldsymbol{x}_1\}$$

is a basis for V if and only if n is odd.

Proof.

It is easy to obtain

$$\begin{pmatrix} \boldsymbol{x}_1 + \boldsymbol{x}_2 & \boldsymbol{x}_2 + \boldsymbol{x}_3 & \cdots & \boldsymbol{x}_{n-1} + \boldsymbol{x}_n & \boldsymbol{x}_n + \boldsymbol{x}_1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_n \end{pmatrix} \boldsymbol{A},$$

where $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 1 \end{pmatrix}$.

($\Rightarrow$) If n is odd, $\det(\boldsymbol{A}) = 1 + (-1)^{1+n} = 2 \neq 0$. Thus T is a basis.

($\Leftarrow$) If n is even, $\det(\boldsymbol{A}) = 1 + (-1)^{1+n} = 2$. Thus T is not a basis.

□

Exercise (Question 1(c) in Final 2008–2009(II))

Give an example of a family of subspaces $V_1, V_2, \ldots, V_n$ *of* $\mathbb{R}^n$ *such that* $\dim(V_i) = i$ *for* $i = 1, 2, \ldots, n$ *and* $V_1 \subset V_2 \subset \cdots \subset V_n$*. Justify your answer.*

Solution.

For any $i = 1, 2, \ldots, n$, let

$$V_i = \{(x_1, x_2, x_3, \ldots, x_n) \mid x_{i+1} = x_{i+2} = \cdots = x_n = 0\}.$$

□

Change log

Last modified: 00:04, March 13, 2011.

Schedule of Tutorial 7

- Any question about last tutorial
- Review concepts: Vector spaces associated with matrices:
  - Row spaces and Column spaces;
  - Rank and Nullity.
- Tutorial: 4.11, 4.16, 4.20, 4.21, 4.23, 4.27
- Additional material:
  - Rank inequalities;
  - 4.7, 4.8, 4.13, 4.17, 4.18, 4.24, 4.25, 4.26;
  - Question 2 in Final of 2001–2002(II);
  - Question 4 in Final of 2005–2006(II);
  - Question 8 in Final of 2006–2007(I);
  - 7a, 7b, 7c, 7d.

Row spaces, Column spaces, and Nullspaces

Def Let $\boldsymbol{A}$ be an $m \times n$ matrix. The row space of $\boldsymbol{A}$ is the subspace of $\mathbb{R}^n$ spanned by the rows of $\boldsymbol{A}$. The column space of $\boldsymbol{A}$ is the subspace of $\mathbb{R}^m$ spanned by the columns of $\boldsymbol{A}$.

- Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be row equivalent matrices, then the row space of $\boldsymbol{A}$ = the row space of $\boldsymbol{B}$.
- Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be row equivalent matrices. Then the following statements hold:
  - A given set of columns of $\boldsymbol{A}$ is linearly independent iff the set of corresponding columns of $\boldsymbol{B}$ is linearly independent;
  - A given set of columns of $\boldsymbol{A}$ forms a basis for the column space of $\boldsymbol{A}$ iff the set of corresponding columns of $\boldsymbol{B}$ forms a basis for the column space of $\boldsymbol{B}$.

Def $\boldsymbol{A} \in \mathbb{R}^{m \times n}$. The solution space of the homogeneous system of linear equations $\boldsymbol{Ax} = \boldsymbol{0}$ is called nullspace of $\boldsymbol{A}$, and $\dim(\text{nullspace of } \boldsymbol{A})$ is called the nullity of $\boldsymbol{A}$, denoted by $\text{nullity}(\boldsymbol{A})$.

- The row spaces of $\boldsymbol{A}$ and $\boldsymbol{B}$ are same iff the nullspaces of $\boldsymbol{A}$ and $\boldsymbol{B}$ are same. (We will prove this result in the Chapter 5.)

Rank

- For simplicity, we use $\mathbb{R}^{m\times n}$ to denote the sets of all $m \times n$ matrices.
- For a matrix $\boldsymbol{A}$, $\dim(\text{row space of } \boldsymbol{A}) = \dim(\text{column space of } \boldsymbol{A})$.

Def The rank of matrix $\boldsymbol{A}$ is the dimension of its row space (or column space), denoted by $\operatorname{rank}(\boldsymbol{A})$.

- If $\boldsymbol{R}$ is a REF of $\boldsymbol{A}$, then

$$\begin{aligned}\operatorname{rank}(\boldsymbol{A}) &= (\#\text{ non-zero rows of } \boldsymbol{R}) = (\#\text{ leading entries of } \boldsymbol{R})\\ &= (\#\text{ pivot columns of } \boldsymbol{R}) = (\#\text{ pivot points of } \boldsymbol{R})\\ &= \text{largest } \# \text{ of linearly independent rows (or columns) in } \boldsymbol{A}\end{aligned}$$

- $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, then $\operatorname{rank}(\boldsymbol{A}) \leq \min\{m, n\}$. If $\operatorname{rank}(\boldsymbol{A}) = \min\{m, n\}$, then $\boldsymbol{A}$ is said to be of full rank.
- $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, then $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}^T)$.
- $\boldsymbol{B}$ is a submatrix of $\boldsymbol{A}$, then $\operatorname{rank}(\boldsymbol{B}) \leq \operatorname{rank}(\boldsymbol{A})$.
- $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, then $\operatorname{rank}(\boldsymbol{A}) + \operatorname{nullity}(\boldsymbol{A}) = (\#\text{ columns of } \boldsymbol{A}) = n$.

Relation between rank and invertibility, rank and consistency

- A square matrix $\boldsymbol{A}$ is of full rank iff $\det(\boldsymbol{A}) \neq 0$.
- Structure Theorem for homogeneous systems: Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, and $\operatorname{rank}(\boldsymbol{A}) = r$. Then
  - if $r = n$, $\boldsymbol{Ax} = \boldsymbol{0}$ has the only trivial solution;
  - if $r < n$, $\boldsymbol{Ax} = \boldsymbol{0}$ has nontrivial solutions, depending on $n - r$ parameters.
- Consistency Theorem for inhomogeneous systems: Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent iff
$$\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A} \mid \boldsymbol{b}).$$
- Structure Theorem for inhomogeneous systems: Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, and $\operatorname{rank}(\boldsymbol{A}) = r$. Assume the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent. Then
  - if $r = n$, $\boldsymbol{Ax} = \boldsymbol{b}$ has unique solution;
  - if $r < n$, the general solution depends on $n - r$ parameters; and a general solution $\boldsymbol{x}$ has the form
$$(\text{a general solution for } \boldsymbol{Ax} = \boldsymbol{0}) + (\text{one particular solution to } \boldsymbol{Ax} = \boldsymbol{b}).$$

Exercise (4.11)

Let $\boldsymbol{A}$ *be the* 3×5 *matrix* $\begin{pmatrix} 1 & -\frac{1}{2} & 0 & 1 & 2 \\ 2 & -1 & 3 & 5 & 7 \\ -4 & 2 & 1 & -3 & -7 \end{pmatrix}$. *Show that* $(-3,0,-1,1,1)^T$, $(-1,2,-1,0,1)^T$, $(0,2,0,-1,1)^T$ *form a basis for the nullspace of* $\boldsymbol{A}$*.*

Proof.

1. It is easy to check that each of the three given vectors satisfy the linear system $\boldsymbol{Ax}=\boldsymbol{0}$. Hence the three vectors are contained in the nullspace of $\boldsymbol{A}$.
2. Applying the working definition, assume $c_1(-3,0,-1,1,1)+c_2(-1,2,-1,0,1)+c_3(0,2,0,-1,1)=(0,0,0,0,0)$. By solving the linear system, we have $c_1=c_2=c_3=0$, hence the three vectors are linearly independent.
3. By Gaussian elimination,

$$\begin{pmatrix} 1 & -\frac{1}{2} & 0 & 1 & 2 \\ 2 & -1 & 3 & 5 & 7 \\ -4 & 2 & 1 & -3 & -7 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -\frac{1}{2} & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

   Thus $\mathrm{rank}(\boldsymbol{A})=2$ and hence $\mathrm{nullity}(\boldsymbol{A})=5-2=3$.

By Theorem 3.5.6, the three vectors forms a basis for the nullspace of $\boldsymbol{A}$. □

Exercise (4.16)

Let $V = \{a(1,2,0,0) + b(0,-1,1,0) + c(0,0,0,1) \mid a, b, c \in \mathbb{R}\}$.

(a) *Find a* 4×4 *matrix* $\boldsymbol{A}$ *such that the row space of* $\boldsymbol{A}$ *is* V.

(b) *Find a* 4×4 *matrix* $\boldsymbol{B}$ *such that the column space of* $\boldsymbol{B}$ *is* V.

(c) *Find a* 4×4 *matrix* $\boldsymbol{C}$ *such that the nullspace of* $\boldsymbol{C}$ *is* V.

Solution.

$(1,2,0,0), (0,-1,1,0), (0,0,0,1)$ are linearly independent, then $\dim(V) = 3$.

(a,b) $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, and $\boldsymbol{B} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$.

(c)

1. The rank of $\boldsymbol{C} = (c_{i,j})_{4\times 4}$ is 1. So we can take the last 3 rows of $\boldsymbol{C}$ to be zero rows. Now it suffices to find $c_{11}, c_{12}, c_{13}, c_{14}$.
2. Since $\boldsymbol{C}(1,2,0,0)^T = \boldsymbol{C}(0,-1,1,0)^T = \boldsymbol{C}(0,0,0,1)^T = \boldsymbol{0}$, then $c_{11} + 2c_{12} = 0$, $c_{12} - c_{13} = 0$, $c_{14} = 0$.
3. Then we can take $\boldsymbol{C}$ to be $\begin{pmatrix} -2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

□

Exercise (4.20)

Suppose $\boldsymbol{A}$ and $\boldsymbol{B}$ are two matrices such that $\boldsymbol{AB} = \boldsymbol{0}$. Show that the column space of $\boldsymbol{B}$ is contained in the nullspace of $\boldsymbol{A}$.

Proof.

1. Let $\boldsymbol{B} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$, where $\boldsymbol{b}_j$ is the jth column of $\boldsymbol{B}$.
2. $$\boldsymbol{AB} = \boldsymbol{0} \Rightarrow (\boldsymbol{A}\boldsymbol{b}_1, \ldots, \boldsymbol{A}\boldsymbol{b}_n) = \boldsymbol{0} \Rightarrow \boldsymbol{A}\boldsymbol{b}_j = \boldsymbol{0} \text{ for all } j,$$
   $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ are contained in the nullspace of $\boldsymbol{A}$.
3. For any element $\boldsymbol{x}$ in the column space of $\boldsymbol{B} = \mathrm{span}\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$, it can be written as $\boldsymbol{x} = c_1\boldsymbol{b}_1 + c_2\boldsymbol{b}_2 + \cdots + c_n\boldsymbol{b}_n$. So we have
   $$\boldsymbol{A}\boldsymbol{x} = c_1\boldsymbol{A}\boldsymbol{b}_1 + c_2\boldsymbol{A}\boldsymbol{b}_2 + \cdots + c_n\boldsymbol{A}\boldsymbol{b}_n = \boldsymbol{0},$$
   that is, $\boldsymbol{x}$ is in the nullspace of $\boldsymbol{A}$.
4. So the column space of $\boldsymbol{B}$ is contained in the nullspace of $\boldsymbol{A}$.

□

Exercise (4.21)

Show that there is no matrix whose row space and nullspace both contain the vector $(1, 1, 1)$.

Proof.

1. Let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_m \end{pmatrix}$ be a matrix where $\boldsymbol{a}_i$ is the ith row of $\boldsymbol{A}$. Let $\boldsymbol{u}$ be any column vector in the nullspace of $\boldsymbol{A}$. Then

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{0} \Rightarrow \begin{pmatrix} \boldsymbol{a}_1\boldsymbol{u} \\ \vdots \\ \boldsymbol{a}_m\boldsymbol{u} \end{pmatrix} = \boldsymbol{0} \Rightarrow \boldsymbol{a}_i\boldsymbol{u} = 0 \text{ for all } i.$$

2. Let $\boldsymbol{b}$ be any vector in the row space of $\boldsymbol{A}$, that is, $\boldsymbol{b} = c_1\boldsymbol{a}_1 + \cdots + c_m\boldsymbol{a}_m$ where $c_1, \ldots, c_m$ are scalars. We have

$$\boldsymbol{b}\boldsymbol{u} = c_1\boldsymbol{a}_1\boldsymbol{u} + \cdots + c_m\boldsymbol{a}_m\boldsymbol{u} = 0.$$

3. Since $\begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \neq 0$, it is impossible to have a matrix whose row space and nullspace both contain the vector $(1, 1, 1)$.

$\square$

Exercise (4.22)

Let $\boldsymbol{A}$ be an $m \times n$ matrix and $\boldsymbol{P}$ an $m \times m$ matrix.

(a) *If $\boldsymbol{P}$ is invertible, show that $\operatorname{rank}(\boldsymbol{PA}) = \operatorname{rank}(\boldsymbol{A})$.*

(b) *Given an example such that $\operatorname{rank}(\boldsymbol{PA}) < \operatorname{rank}(\boldsymbol{A})$.*

(c) *Suppose $\operatorname{rank}(\boldsymbol{PA}) = \operatorname{rank}(\boldsymbol{A})$. Is it true that $\boldsymbol{P}$ must be invertible? Justify your answer.*

Proof and Solution.

(a) $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{P}^{-1}\boldsymbol{PA}) \leq \operatorname{rank}(\boldsymbol{PA}) \leq \operatorname{rank}(\boldsymbol{A})$.

(b) $\boldsymbol{P} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, $\boldsymbol{A} = \boldsymbol{I}_2$, then $\operatorname{rank}(\boldsymbol{PA}) = 0 \neq 2 = \operatorname{rank}(\boldsymbol{A})$.

(c) No. For example, let $\boldsymbol{P} = \boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, then $\operatorname{rank}(\boldsymbol{PA}) = 1 = \operatorname{rank}(\boldsymbol{A})$.

$\square$

Exercise (4.23(a))

Let $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *be* $m \times p$ *and* $p \times n$ *matrices respectively. Show that the nullspace of* $\boldsymbol{B}$ *is a subset of the nullspace of* $\boldsymbol{AB}$*. Hence prove that* $\operatorname{rank}(\boldsymbol{AB}) \leq \operatorname{rank}(\boldsymbol{B})$*.*

Proof.

1. Let $\boldsymbol{u}$ be any vector in the nullspace of $\boldsymbol{B}$, that is, $\boldsymbol{Bu} = \boldsymbol{0}$.
2. Since $\boldsymbol{ABu} = \boldsymbol{A0} = \boldsymbol{0}$, $\boldsymbol{u}$ is a vector in the nullspace of $\boldsymbol{AB}$.
3. So the nullspace of $\boldsymbol{B}$ is a subset of the nullspace of $\boldsymbol{AB}$, and hence $\operatorname{nullity}(\boldsymbol{B}) \leq \operatorname{nullity}(\boldsymbol{AB})$.
4. Therefore
$$\operatorname{rank}(\boldsymbol{AB}) = n - \operatorname{nullity}(\boldsymbol{AB}) \leq n - \operatorname{nullity}(\boldsymbol{B}) = \operatorname{rank}(\boldsymbol{B}).$$
5. $\operatorname{rank}(\boldsymbol{B}^T\boldsymbol{A}^T) \leq \operatorname{rank}(\boldsymbol{A}^T)$. Since $\operatorname{rank}(\boldsymbol{A}^T) = \operatorname{rank}(\boldsymbol{A})$ and $\operatorname{rank}(\boldsymbol{AB}) = \operatorname{rank}(\boldsymbol{B}^T\boldsymbol{A}^T)$, $\operatorname{rank}(\boldsymbol{AB}) \leq \operatorname{rank}(\boldsymbol{A})$.

□

Exercise (4.23(b))

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be $m \times p$ and $p \times n$ matrices respectively. Show that every column of the matrix $\boldsymbol{AB}$ lies in the column space of $\boldsymbol{A}$. Hence, or otherwise, prove that $\operatorname{rank}(\boldsymbol{AB}) \leq \operatorname{rank}(\boldsymbol{A})$.

Proof.

1. Let $\boldsymbol{x}_j$ be the jth column of $\boldsymbol{AB}$. Then $\boldsymbol{x}_j = \boldsymbol{A}\boldsymbol{b}_j$ where $\boldsymbol{b}_j$ is the jth column of $\boldsymbol{B}$.
2. Let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_p \end{pmatrix}$, where $\boldsymbol{a}_i$ is the ith column of $\boldsymbol{A}$, and let $\boldsymbol{b}_j = (b_{1j}, b_{2j}, \ldots, b_{pj})^T$. Then

$$\boldsymbol{x}_j = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_p \end{pmatrix} \begin{pmatrix} b_{1j} \\ \vdots \\ b_{pj} \end{pmatrix} = b_{1j}\boldsymbol{a}_1 + b_{2j}\boldsymbol{a}_2 + \cdots + b_{pj}\boldsymbol{a}_p.$$

3. Hence, $\boldsymbol{x}_j$ is in the column space of $\boldsymbol{A}$.
4. Therefore the column space of $\boldsymbol{AB}$ is contained in the column space of $\boldsymbol{A}$, and hence

$$\begin{aligned} \operatorname{rank}(\boldsymbol{AB}) &= \dim(\text{the column space of } \boldsymbol{AB}) \\ &\leq \dim(\text{the column space of } \boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}) \end{aligned}$$

$\square$

Exercise (4.27)

Determine which of the following statements are true. Justify your answer.

(a) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are two row equivalent matrices, then the row space of $\boldsymbol{A}^T$ and the row space of $\boldsymbol{B}^T$ are the same.*

(b) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are two row equivalent matrices, then the column space of $\boldsymbol{A}^T$ and the column space of $\boldsymbol{B}^T$ are the same.*

(c) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are two row equivalent matrices, then the nullspace of $\boldsymbol{A}^T$ and the nullspace of $\boldsymbol{B}^T$ are the same.*

(d) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are two matrices of the same size, then $\mathrm{rank}(\boldsymbol{A}+\boldsymbol{B}) = \mathrm{rank}(\boldsymbol{A}) + \mathrm{rank}(\boldsymbol{B})$.*

(e) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are two matrices of the same size, then $\mathrm{nullity}(\boldsymbol{A}+\boldsymbol{B}) = \mathrm{nullity}(\boldsymbol{A}) + \mathrm{nullity}(\boldsymbol{B})$.*

(f) *If $\boldsymbol{A}$ is an $n \times m$ matrix and $\boldsymbol{B}$ is an $m \times n$ matrix, then $\mathrm{rank}(\boldsymbol{AB}) = \mathrm{rank}(\boldsymbol{BA})$.*

(g) *If $\boldsymbol{A}$ is an $n \times m$ matrix and $\boldsymbol{B}$ is an $m \times n$ matrix, then $\mathrm{nullity}(\boldsymbol{AB}) = \mathrm{nullity}(\boldsymbol{BA})$.*

Solution.

(ac) False. For example, let $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\boldsymbol{B} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

(b) True. Since the row space of $\boldsymbol{A}$ and the the row space of $\boldsymbol{B}$ are the same. Hence the column space of $\boldsymbol{A}^T$ and the column space of $\boldsymbol{B}^T$ are the same.

(d) False. For example, let $\boldsymbol{A} = \boldsymbol{B} = \boldsymbol{I}_1$.

(e) False. For example, let $\boldsymbol{A} = \boldsymbol{B} = \mathbf{0}_1$.

(fg) False. For example, let $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\boldsymbol{B} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

$\square$

(1) $\operatorname{rank}(\boldsymbol{AB}) \leq \min\{\operatorname{rank}(\boldsymbol{A}), \operatorname{rank}(\boldsymbol{B})\}$. See Exercise 4.23.

(2) $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{P} \in \mathbb{R}^{m\times m}$ and $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ are invertible, then $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{PA}) = \operatorname{rank}(\boldsymbol{AQ}) = \operatorname{rank}(\boldsymbol{PAQ})$. By (1) or see Exercise 4.22.

(2a) $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\operatorname{rank}(\boldsymbol{A}) = r \leq \min\{m, n\}$, then there exist invertible matrices $\boldsymbol{P} \in \mathbb{R}^{m\times m}$ and $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$, such that $\boldsymbol{PAQ} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0}_{(n-r)\times(m-r)} \end{pmatrix}$. By (2).

(2b) $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C} \end{pmatrix}$, then $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{B}) + \operatorname{rank}(\boldsymbol{C})$. By (2a).

(2c) $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\operatorname{rank}(\boldsymbol{A}) = r$, then there exist $\boldsymbol{B} \in \mathbb{R}^{m\times r}$ and $\boldsymbol{C} \in \mathbb{R}^{r\times n}$, such that $\boldsymbol{A} = \boldsymbol{BC}$. By (2a).

(3) $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{p\times q}$, $\boldsymbol{C} \in \mathbb{R}^{m\times p}$, then $\operatorname{rank}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix} \geq \operatorname{rank}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix}$. By (2).

(3a) Sylvester's inequality: $\boldsymbol{A} \in \mathbb{R}^{m\times p}$, $\boldsymbol{B} \in \mathbb{R}^{p\times n}$, then $\operatorname{rank}(\boldsymbol{AB}) \geq \operatorname{rank}(\boldsymbol{A}) + \operatorname{rank}(\boldsymbol{B}) - p$. By (2), (3).

(3b) Frobenius's inequality: $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times p}$, $\boldsymbol{C} \in \mathbb{R}^{p\times q}$, then $\operatorname{rank}(\boldsymbol{AB}) + \operatorname{rank}(\boldsymbol{BC}) - \operatorname{rank}(\boldsymbol{B}) \leq \operatorname{rank}(\boldsymbol{ABC})$. By (2), (3).

(4) $\operatorname{rank}(\boldsymbol{A} \pm \boldsymbol{B}) \leq \operatorname{rank}\left(\boldsymbol{A} \mid \boldsymbol{B}\right) \leq \operatorname{rank}(\boldsymbol{A}) + \operatorname{rank}(\boldsymbol{B})$. By (2), (2b) and Def.

(5) $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, then $\operatorname{rank}(\boldsymbol{AA}^T) = \operatorname{rank}(\boldsymbol{A}^T\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A})$. Def.

(2a) Apply elementary row operations, we can get a RREF; then apply elementary column operations, we can get the form $\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$, where $r = (\#\text{ leading entries}) = \operatorname{rank}(\boldsymbol{A})$.

(2b) By (2a), there exist invertible matrices $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$, such that

$$\boldsymbol{P}_1\boldsymbol{B}\boldsymbol{Q}_1 = \begin{pmatrix} \boldsymbol{I}_{r_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}, \quad \boldsymbol{P}_2\boldsymbol{C}\boldsymbol{Q}_2 = \begin{pmatrix} \boldsymbol{I}_{r_2} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

where $r_1 = \operatorname{rank}(\boldsymbol{B})$ and $r_2 = \operatorname{rank}(\boldsymbol{C})$. Then

$$\begin{pmatrix} \boldsymbol{P}_1 & \\ & \boldsymbol{P}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{B} & \\ & \boldsymbol{C} \end{pmatrix} \begin{pmatrix} \boldsymbol{Q}_1 & \\ & \boldsymbol{Q}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{I}_{r_1} & & & \\ & \boldsymbol{0} & & \\ & & \boldsymbol{I}_{r_2} & \\ & & & \boldsymbol{0} \end{pmatrix}.$$

Therefore $\operatorname{rank}(\boldsymbol{A}) = r_1 + r_2 = \operatorname{rank}(\boldsymbol{B}) + \operatorname{rank}(\boldsymbol{C})$.

(2c) By (2a), we have $\boldsymbol{A} = \boldsymbol{P}^{-1} \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \boldsymbol{Q}^{-1} = \boldsymbol{P}^{-1} \begin{pmatrix} \boldsymbol{I}_r \\ \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \end{pmatrix} \boldsymbol{Q}^{-1}$. Let $\boldsymbol{B} = \boldsymbol{P}^{-1} \begin{pmatrix} \boldsymbol{I}_r \\ \boldsymbol{0} \end{pmatrix}$ and $\boldsymbol{C} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \end{pmatrix} \boldsymbol{Q}^{-1}$, then $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}$.

(3) By (2a), there exist invertible matrices $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$, such that

$$\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1 = \begin{pmatrix} \boldsymbol{I}_{r_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}, \quad \boldsymbol{P}_2\boldsymbol{B}\boldsymbol{Q}_2 = \begin{pmatrix} \boldsymbol{I}_{r_2} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

where $r_1 = \operatorname{rank}(\boldsymbol{B})$ and $r_2 = \operatorname{rank}(\boldsymbol{C})$. Then

$$\begin{pmatrix} \boldsymbol{P}_1 & \\ & \boldsymbol{P}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ & \boldsymbol{B} \end{pmatrix} \begin{pmatrix} \boldsymbol{Q}_1 & \\ & \boldsymbol{Q}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{I}_{r_1} & & \boldsymbol{P}_1\boldsymbol{C}\boldsymbol{Q}_2 & \\ & \boldsymbol{0} & & \\ & & \boldsymbol{I}_{r_2} & \\ & & & \boldsymbol{0} \end{pmatrix}.$$

Therefore $\operatorname{rank}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix} \geq \operatorname{rank}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix}$.

(3a)

$$\begin{pmatrix} \boldsymbol{I}_m & \boldsymbol{A} \\ & \boldsymbol{I}_p \end{pmatrix} \begin{pmatrix} \boldsymbol{AB} & \\ & \boldsymbol{I}_p \end{pmatrix} \begin{pmatrix} \boldsymbol{I}_n & \\ -\boldsymbol{B} & \boldsymbol{I}_p \end{pmatrix} \begin{pmatrix} & -\boldsymbol{I}_p \\ \boldsymbol{I}_n & \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} & \\ \boldsymbol{I} & \boldsymbol{B} \end{pmatrix}.$$

Hence by (3) $\operatorname{rank}(\boldsymbol{AB}) + p \geq \operatorname{rank}(\boldsymbol{A}) + \operatorname{rank}(\boldsymbol{B})$.

(3b)

$$\begin{pmatrix} \boldsymbol{I}_m & \boldsymbol{A} \\ & \boldsymbol{I}_n \end{pmatrix} \begin{pmatrix} \boldsymbol{ABC} & \\ & \boldsymbol{B} \end{pmatrix} \begin{pmatrix} \boldsymbol{I}_q & \\ -\boldsymbol{C} & \boldsymbol{I}_p \end{pmatrix} \begin{pmatrix} & -\boldsymbol{I}_q \\ \boldsymbol{I}_p & \end{pmatrix} = \begin{pmatrix} \boldsymbol{AB} & \\ \boldsymbol{B} & \boldsymbol{BC} \end{pmatrix}.$$

Hence by (3) $\mathrm{rank}(\boldsymbol{ABC}) + \mathrm{rank}(\boldsymbol{B}) \geq \mathrm{rank}(\boldsymbol{AB}) + \mathrm{rank}(\boldsymbol{BC})$.

(4)

$$\mathrm{rank}(\boldsymbol{A}+\boldsymbol{B}) \leq \mathrm{rank} \begin{pmatrix} \boldsymbol{A}+\boldsymbol{B} & \\ & \boldsymbol{0} \end{pmatrix} = \mathrm{rank} \begin{pmatrix} \boldsymbol{A}+\boldsymbol{B} & \boldsymbol{B} \\ & \boldsymbol{0} \end{pmatrix} = \mathrm{rank}(\boldsymbol{A} \mid \boldsymbol{B}).$$

largest $\#$ l.i. columns in $(\boldsymbol{A} \mid \boldsymbol{B}) \leq$
largest $\#$ l.i. columns in $\boldsymbol{A}$ + largest $\#$ l.i. columns in $\boldsymbol{B}$, so
$\mathrm{rank}(\boldsymbol{A} \mid \boldsymbol{B}) \leq \mathrm{rank}(\boldsymbol{A}) + \mathrm{rank}(\boldsymbol{B})$.

Exercise (4.7)

Let $V = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\}$ *where*

$$\boldsymbol{u}_1 = (1,1,1,1,1),\ \boldsymbol{u}_2 = (1,x,x,x,x),\ \boldsymbol{u}_3 = (1,x,x^2,x,x^2),\ \boldsymbol{u}_4 = (1,x^3,x,2x-x^3,x)$$

for some constant x. *Find a basis for* V *and determine the dimension of* V.

Solution.

By Gaussian elimination, we have

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

- If $x = 1$, then $\{\boldsymbol{u}_1\}$ is a basis for V and $\dim(V) = 1$.
- If $x = 0$, then $\{\boldsymbol{u}_1, (0,1,1,1,1)\}$ is a basis for V and $\dim(V) = 2$.
- If $x = -1$, then $\{\boldsymbol{u}_1, (0,-2,-2,-2,-2), (0,0,2,0,2)\}$ is a basis for V and $\dim(V) = 3$.
- If $x \notin \{0,1,-1\}$, then $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\}$ is a basis for V and $\dim(V) = 4$.

□

Exercise (4.8)

For each of the following cases, write down a matrix with the required property or explain why no such matrix exists.

(a) *Column space contains vectors* $(1,0,0)^T, (0,0,1)^T$ *and row space contains vectors* $(1,1), (1,2)$.

(b) *Column space* $= \mathbb{R}^4$, *row space* $= \mathbb{R}^3$.

Solution.

(a) Yes, for example: $\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$.

(b) No. By Theorem 4.2.1, the dimensions of the row space and column space of a matrix must be the same.

□

Exercise (4.13)

Determine the possible rank and nullity of each of the following matrices:

$$\text{(a) } \boldsymbol{A} = \begin{pmatrix} 1 & 1 & a \\ 1 & a & 1 \\ a & 1 & 1 \end{pmatrix}, \quad \text{(b) } \boldsymbol{B} = \begin{pmatrix} 0 & 0 & b \\ 0 & 0 & c \\ d & e & f \end{pmatrix},$$

where a, b, c, d, e, f *are real numbers.*

Solution of part (a).

By Gaussian elimination:

$$\begin{pmatrix} 1 & 1 & a \\ 1 & a & 1 \\ a & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & a \\ 0 & a-1 & 1-a \\ 0 & 1-a & 1-a^2 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & a \\ 0 & a-1 & 1-a \\ 0 & 0 & -(a-1)(a+2) \end{pmatrix}.$$

- when $a = 1$, there is only 1 non-zero row, that is, $\text{rank}(\boldsymbol{A}) = 1$, $\text{nullity}(\boldsymbol{A}) = 2$;
- when $a = -2$, there are 2 non-zero rows, that is, $\text{rank}(\boldsymbol{A}) = 2$, $\text{nullity}(\boldsymbol{A}) = 1$;
- For other cases, all of the rows are non-zero rows, that is, $\text{rank}(\boldsymbol{A}) = 3$, $\text{nullity}(\boldsymbol{A}) = 0$.

$\square$

Solution of part (b).

For

$$\boldsymbol{B} = \begin{pmatrix} 0 & 0 & b \\ 0 & 0 & c \\ d & e & f \end{pmatrix},$$

1. the first 2 rows are linearly dependent, then $\mathrm{rank}(\boldsymbol{B}) \leq 2$.
2. if $b = c = d = e = f = 0$, $\mathrm{rank}(\boldsymbol{B}) = 0, \mathrm{nullity}(\boldsymbol{B}) = 3$;
3. if either (i) $b = c = 0$ and not all d, e, f are zero or (ii) $d = e = 0$ and not all b, c, f are zero, $\mathrm{rank}(\boldsymbol{B}) = 1, \mathrm{nullity}(\boldsymbol{B}) = 2$.
4. if not all b, c are zero and not all d, e are zero, $\mathrm{rank}(\boldsymbol{B}) = 2, \mathrm{nullity}(\boldsymbol{B}) = 1$.

$\square$

Exercise (4.17)

Let $\boldsymbol{A}$ be a 3×4 matrix. Suppose that $x_1 = 1$, $x_2 = 0$, $x_3 = -1$, $x_4 = 0$ is a solution to a non-homogeneous linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ and that the homogeneous system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ has a general solution $x_1 = t - 2s$, $x_2 = s + t$, $x_3 = s$, $x_4 = t$ where s, t are arbitrary parameters.

(a) *Find a basis for the nullspace of $\boldsymbol{A}$ and determine the nullity of $\boldsymbol{A}$.*

(b) *Find a general solution for the system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.*

(c) *Write down the RREF of $\boldsymbol{A}$.*

(d) *Find a basis for the row space of $\boldsymbol{A}$ and determine the rank of $\boldsymbol{A}$.*

(e) *Do we have enough information for us to find the column space of $\boldsymbol{A}$?*

Solution of parts (a,b).

(a) Since $(x_1, x_2, x_3, x_4)^T = (t - 2s, s + t, s, t)^T = s(-2, 1, 1, 0)^T + t(1, 1, 0, 1)^T$, $\{(-2, 1, 1, 0)^T, (1, 1, 0, 1)^T\}$ is a basis for the nullspace of $\boldsymbol{A}$. The nullity of $\boldsymbol{A}$ is 2.

(b) A general solution of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is $x_1 = t - 2s + 1$, $x_2 = s + t$, $x_3 = s - 1$, $x_4 = t$ where s, t are arbitrary parameters.

$\square$

Solution of parts (c–e).

(c)
- It is obvious that $\mathrm{nullity}(\boldsymbol{A}) = 2$, and $\mathrm{rank}(\boldsymbol{A}) = 1$. So we have that the last row in the RREF of $\boldsymbol{A}$ is a zero row.
- A general solution of $\boldsymbol{Ax} = \boldsymbol{0}$ is $\begin{cases} x_1 = -2s + t \\ x_2 = s + t \\ x_3 = s \\ x_4 = t \end{cases}$. Now we want to find 2 (since $\mathrm{rank}(\boldsymbol{A}) = 2$) equations for x_1, x_2, x_3, x_4: $\begin{cases} x_1 = -2x_3 + x_4 \\ x_2 = x_3 + x_4 \end{cases}$.
- Hence, the entries in the i-th row of RREF are the coefficients in the i-th condition $(i = 1, 2)$, that is, RREF is $\begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

(d) $\{(1, 0, 2, -1), (0, 1, -1, -1)\}$ is a basis for the row space of $\boldsymbol{A}$. The rank of $\boldsymbol{A}$ is 2.

(e) No, we cannot find the column space of $\boldsymbol{A}$ with the given information.

$\square$

Exercise (4.18)

Let $A = (a_1 \, a_2 \, a_3 \, a_4 \, a_5)$ be a 4×5 matrix such that the columns a_1, a_2, a_3 are linearly independent while $a_4 = a_1 - 2a_2 + a_3$ and $a_5 = a_2 + a_3$.

(a) *Determine the RREF of $\boldsymbol{A}$.*

(b) *Find a basis for the row space of $\boldsymbol{A}$ and a basis for the column space of $\boldsymbol{A}$.*

Solution.

(a) Let $\boldsymbol{R}$ be the RREF of $\boldsymbol{A}$. Since a_1, a_2, a_3 are linearly independent, the first three columns of $\boldsymbol{R}$ are linearly independent. Thus the first three columns of $\boldsymbol{R}$ must be $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Since $\begin{cases} a_4 = a_1 - 2a_2 + a_3 \\ a_5 = a_2 + a_3 \end{cases}$, $\boldsymbol{R} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -2 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

(b) It is obvious that $\{a_1, a_2, a_3\}$ is a basis for the column space of $\boldsymbol{A}$, and both the dimensions of column space and row spaces are 3. Hence $\{(1,0,0,1,0), (0,1,0,-2,1), (0,0,1,1,1)\}$ is a basis for the row space of $\boldsymbol{A}$.

$\square$

Exercise (4.24)

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two matrices of the same size. Show that

$$\operatorname{rank}(\boldsymbol{A}+\boldsymbol{B}) \leq \operatorname{rank}(\boldsymbol{A}) + \operatorname{rank}(\boldsymbol{B}).$$

Proof.

$$\begin{aligned}\operatorname{rank}(\boldsymbol{A}) + \operatorname{rank}(\boldsymbol{B}) &= \operatorname{rank}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{B}\end{pmatrix} = \operatorname{rank}\begin{pmatrix}\boldsymbol{I} & \boldsymbol{I}\\ \boldsymbol{0} & \boldsymbol{I}\end{pmatrix}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{B}\end{pmatrix}\begin{pmatrix}\boldsymbol{I} & \boldsymbol{I}\\ \boldsymbol{0} & \boldsymbol{I}\end{pmatrix}\\ &= \operatorname{rank}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{A}+\boldsymbol{B}\\ \boldsymbol{0} & \boldsymbol{B}\end{pmatrix}\\ &\geq \operatorname{rank}\begin{pmatrix}\boldsymbol{0} & \boldsymbol{A}+\boldsymbol{B}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix} = \operatorname{rank}(\boldsymbol{A}+\boldsymbol{B}).\end{aligned}$$

□

Exercise (4.25)

Let $\boldsymbol{A}$ be an $m \times n$ matrix.

(a) *Show that the nullspace of $\boldsymbol{A}$ is equal to the nullspace of $\boldsymbol{A}^T\boldsymbol{A}$.*

(b) *Show that* $\mathrm{nullity}(\boldsymbol{A}) = \mathrm{nullity}(\boldsymbol{A}^T\boldsymbol{A})$ *and* $\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A}^T\boldsymbol{A})$.

(c) *Is it true that* $\mathrm{nullity}(\boldsymbol{A}) = \mathrm{nullity}(\boldsymbol{A}\boldsymbol{A}^T)$? *Justify your answer.*

(d) *Is it true that* $\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A}\boldsymbol{A}^T)$? *Justify your answer.*

Proof and Solution.

(a) Proved in lecture;

(b) By part (a);

(c) No. For example, $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

(d) Yes. By (b), $\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A}^T) = \mathrm{rank}((\boldsymbol{A}^T)^T\boldsymbol{A}^T) = \mathrm{rank}(\boldsymbol{A}\boldsymbol{A}^T)$.

□

Exercise (4.26)

Let $\boldsymbol{A}$ be an $m \times n$ matrix. Suppose the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent for any $\boldsymbol{b} \in \mathbb{R}^m$. Show that the linear system $\boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0}$ has only the trivial solution.

Proof.

- $\boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0} \Rightarrow \boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0} \Rightarrow \boldsymbol{b}^T\boldsymbol{y} = \boldsymbol{0}$ for any $\boldsymbol{b} \in \mathbb{R}^m$.
- For any $1 \leq i \leq m$, $\boldsymbol{b} = \boldsymbol{e}_i$ whose components are zeros except i-th component, then i-th component of $\boldsymbol{y}$ is 0, that is, $\boldsymbol{y} = \boldsymbol{0}$.

$\square$

Exercise (Question 2 in Final of 2001-2002(II), Question 4 in Final of 2005-2006(II))

Determine the possible rank of each of the following matrices:

$$\begin{pmatrix} 1 & 1 & x^2 \\ 1 & x^2 & 1 \\ x^2 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{pmatrix},$$

where x, a, b, c *are real numbers.*

Exercise (Question 8 in Final of 2006-2007(I))

(a) *Let* $\boldsymbol{A}$ *be a square matrix such that* $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}^2)$.

  (i) *Show that the nullspace of* $\boldsymbol{A}$ *is equal to the nullspace of* $\boldsymbol{A}^2$.

  (ii) *Show that the nullspace of* $\boldsymbol{A}$ *and the column space of* $\boldsymbol{A}$ *intersect trivially.*

(b) *Suppose there exist* $n \times n$ *matrices* $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$ *such that* $\boldsymbol{X}\boldsymbol{Y} = \boldsymbol{Z}$. *Show that the column space of* $\boldsymbol{Z}$ *is a subset of the column space of* $\boldsymbol{X}$.

(c) *Let* $\boldsymbol{B} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

  (i) *Find the nullspace of* $\boldsymbol{B}^2$.

  (ii) *Show that there does not exist any* 3×3 *matrix* $\boldsymbol{C}$ *such that* $\boldsymbol{C}^2 = \boldsymbol{B}$.

Exercise (7a)

$\boldsymbol{A} \in \mathbb{R}^{n\times n}$, *then*

(a) $\operatorname{rank}(\operatorname{adj}(\boldsymbol{A})) = n$ *iff* $\operatorname{rank}(\boldsymbol{A}) = n$;

(b) $\operatorname{rank}(\operatorname{adj}(\boldsymbol{A})) = 1$ *iff* $\operatorname{rank}(\boldsymbol{A}) = n-1$;

(c) $\operatorname{rank}(\operatorname{adj}(\boldsymbol{A})) = 0$ *iff* $\operatorname{rank}(\boldsymbol{A}) < n-1$;

Proof.

(a) $\boldsymbol{A}$ is invertible iff $\operatorname{adj}(\boldsymbol{A})$ is invertible.

(b) $\operatorname{rank}(\boldsymbol{A}) = n-1$ iff $\boldsymbol{A} = \boldsymbol{P}\begin{pmatrix} \boldsymbol{I}_{n-1} & \\ & 0 \end{pmatrix}\boldsymbol{Q}$ iff

$\operatorname{adj}(\boldsymbol{A}) = \operatorname{adj}(\boldsymbol{Q})\begin{pmatrix} \boldsymbol{0}_{n-1} & \\ & 1 \end{pmatrix}\operatorname{adj}(\boldsymbol{P})$.

(c) $\operatorname{rank}(\boldsymbol{A}) < n-1$ iff $\operatorname{adj}(\boldsymbol{A}) = \boldsymbol{0}$ iff $\operatorname{rank}(\operatorname{adj}(\boldsymbol{A})) = 0$.

$\square$

Exercise (7b)

$\boldsymbol{A} \in \mathbb{R}^{n\times n}$, *and* $\boldsymbol{A}^2 = \boldsymbol{A}$, *then* $\operatorname{rank}(\boldsymbol{A}) = \operatorname{tr}(\boldsymbol{A})$.

Proof.

Let $\operatorname{rank}(\boldsymbol{A}) = r$, then there exist invertible matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$, such that $\boldsymbol{A} = \boldsymbol{P}\begin{pmatrix} \boldsymbol{I}_r & \\ & \boldsymbol{0} \end{pmatrix}\boldsymbol{Q}$. Since $\boldsymbol{A}^2 = \boldsymbol{A}$, we have $\boldsymbol{A} = \boldsymbol{P}\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{R}_{12} \\ & \boldsymbol{0} \end{pmatrix}\boldsymbol{P}^{-1}$. Hence $\operatorname{tr}(\boldsymbol{A}) = \operatorname{tr}\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{R}_{12} \\ & \boldsymbol{0} \end{pmatrix} = r = \operatorname{rank}(\boldsymbol{A})$. □

Exercise (7c)

$\boldsymbol{A} \in \mathbb{R}^{n\times n}$,

(a) *if there exists an integer* k, *such that* $\operatorname{rank}(\boldsymbol{A}^k) = \operatorname{rank}(\boldsymbol{A}^{k+1})$, *then* $\operatorname{rank}(\boldsymbol{A}^k) = \operatorname{rank}(\boldsymbol{A}^{k+1}) = \operatorname{rank}(\boldsymbol{A}^{k+2}) = \cdots$.

(b) *there exists an integer* k, *such that* $\operatorname{rank}(\boldsymbol{A}^k) = \operatorname{rank}(\boldsymbol{A}^{k+1})$.

Proof.

By Frobenius's inequality. □

Exercise (7d)

$\boldsymbol{A} \in \mathbb{R}^{n\times n}$, *does* $\operatorname{rank}(\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^T) = \operatorname{rank}(\boldsymbol{I} - \boldsymbol{A}^T\boldsymbol{A})$ *hold?*

Proof.

By the following equations:

$$\begin{pmatrix} \boldsymbol{I} & -\boldsymbol{A} \\ & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{I} \end{pmatrix} = \begin{pmatrix} \boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^T & \\ \boldsymbol{A}^T & \boldsymbol{I} \end{pmatrix},$$

$$\begin{pmatrix} \boldsymbol{I} & \\ -\boldsymbol{A}^T & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{I} \end{pmatrix} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{A} \\ & \boldsymbol{I} - \boldsymbol{A}^T\boldsymbol{A} \end{pmatrix}.$$

$\square$

Change log

- Page 153: Add a remark for “the relation between nullspace and row space”;
- Page 159: Revise a typo: “$\mathbf{0}$” to “0”.

Last modified: 23:38, March 19, 2011.

Schedule of Tutorial 8

- Any question about last tutorial
- Review concepts:
  - Eigenvalue, Eigenvector and Eigenspace;
  - Diagonalization.
- Tutorial: 6.6, 6.10, 6.13, 6.14, 6.16, 6.18
- Additional material:
  - The algebraic multiplicity, the geometric multiplicity;
  - Remak 6.2.5.2 and Remak 6.2.5.3;
  - Exercise 6.3, 6.7, 6.12;
  - Question 6(b) in Final of 2006–2007(II);
  - Question 5 in Final of 2004–2005(II);
  - Question 1(a) in Final of 2005–2006(I);
  - Question 4(b) in Final of 2006-02007(II);
  - Question 3(b-iii) in Final of 2009–2010(I).

Eigenvalue and Eigenvector

Here we focus on the real case. Let $\boldsymbol{A}$ be a real square matrix of order n.

- If there exist a nonzero column vector $\boldsymbol{x} \in \mathbb{R}^n$ and a (real) scalar λ such that $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, then λ is called an eigenvalue of $\boldsymbol{A}$, and $\boldsymbol{x}$ is said to be an eigenvector of $\boldsymbol{A}$ associated with the eigenvalue λ.
- The equation $\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$ is called the characteristic equation of $\boldsymbol{A}$ and the polynomial $\varphi(\lambda) = \det(\lambda\boldsymbol{I} - \boldsymbol{A})$ is called the characteristic polynomial of $\boldsymbol{A}$.
- λ is an eigenvalue iff $\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$. Hence, (# eigenvalues) $\leq n$.
- If $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, where $\boldsymbol{P}$ is an invertible matrix, then $\boldsymbol{A}$ and $\boldsymbol{B}$ have same eigenvalues. While the converse is not necessarily true. (See Exercise 6.13)
- λ_1 and λ_2 are 2 distinct eigenvalues, $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are 2 eigenvectors associated with λ_1 and λ_2, respectively. Then $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are linearly independent.
- If λ is an eigenvalue of $\boldsymbol{A}$, then $c\lambda$ is an eigenvalue of $c\boldsymbol{A}$. (See Exercise 6.6(c))
- If $\boldsymbol{A}$ has n eigenvalues $\{\lambda_i\}_{i=1}^n$, then $\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^n \lambda_i$, $\det(\boldsymbol{A}) = \prod_{i=1}^n \lambda_i$. (See Exercise 6.2(a))
- $\boldsymbol{A}\boldsymbol{B}$ and $\boldsymbol{B}\boldsymbol{A}$ have same eigenvalues.
- Cayley-Hamilton's Theorem: If $\varphi(\lambda)$ is the characteristic polynomial, then $\varphi(\boldsymbol{A}) = \boldsymbol{0}$. (See Exercise 6.2(b))

Algebraic multiplicity and Geometric multiplicity

Let $\boldsymbol{A}$ be a real square matrix of order n. Then the characteristic polynomial $\varphi_{\boldsymbol{A}}(\lambda)$ can be decomposed as

$$(\lambda - \lambda_1)^{r_1} \cdots (\lambda - \lambda_k)^{r_k} (\lambda^2 + a_1 \lambda + b_1)^{s_1} \cdots (\lambda^2 + a_l \lambda + b_l)^{s_l}.$$

(See Remark 6.2.5.1)

- Let λ be an eigenvalue of $\boldsymbol{A}$. Then the solution space of the linear system $(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = 0$ is called the eigenspace of $\boldsymbol{A}$ associated with the eigenvalue λ and is denoted by $E_\lambda = \{\boldsymbol{x} \in \mathbb{R}^n \mid (\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}\}$.
- The geometric multiplicity of an eigenvalue is defined as the dimension of the associated eigenspace.
- The algebraic multiplicity of an eigenvalue is defined as the multiplicity of the corresponding root of the characteristic polynomial. That is, the algebraic multiplicity of λ_i is r_i for $i = 1, 2, \ldots, k$.
- For any eigenvalue λ of $\boldsymbol{A}$,

$$(\text{the algebraic multiplicity of } \lambda) \geq (\text{ the geometric multiplicity of } \lambda) \geq 1.$$

Diagonalization

Let $\boldsymbol{A}$ be a real square matrix of order n.

- $\boldsymbol{A}$ is called diagonalizable if there exists an invertible matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is a diagonal matrix.
- $\boldsymbol{A}$ is diagonalizable iff $\boldsymbol{A}$ has n linearly independent eigenvectors.
- If $\boldsymbol{A}$ has n distinct eigenvalues, then $\boldsymbol{A}$ is diagonalizable; while the converse is not necessarily true. That is, if $\boldsymbol{A}$ is diagonalizable, $\boldsymbol{A}$ may have some same eigenvalues (e.g. $\boldsymbol{I}_2$).
- $\boldsymbol{A}$ is diagonalizable iff for each eigenvalue λ_0 of matrix $\boldsymbol{A}$, the algebraic multiplicity is equal to the geometric multiplicity.
- Schur's Theorem: There exists an invertible matrix $\boldsymbol{P}$, such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is an upper-triangular block matrix.

How To

How to determine whether a square matrix is diagonalizable?

- Method 1:
  (1) Solve $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$ to find all distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$.
  (2) For each eigenvalue λ_i, find a basis S_{λ_i} for the eigenspace E_{λ_i}.
  (3) Let $S = S_{\lambda_1} \cup S_{\lambda_2} \cup \cdots S_{\lambda_k}$.
    (a) If $|S| < n$, then $\boldsymbol{A}$ is not diagonalizable.
    (b) If $|S| = n$, say $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$, then the square matrix $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n \end{pmatrix}$ diagonalizes $\boldsymbol{A}$.

- Method 2:
  (1) Decompose the characteristic polynomial $\varphi_{\boldsymbol{A}}(\lambda)$ as
  $$(\lambda - \lambda_1)^{r_1} \cdots (\lambda - \lambda_k)^{r_k} (\lambda^2 + a_1\lambda + b_1)^{s_1} \cdots (\lambda^2 + a_l\lambda + b_l)^{s_l},$$
  where $\lambda_1, \ldots, \lambda_k$ are pairwise distinct, $(\lambda^2 + a_j\lambda + b_j)$ can not do more decomposition.
    (a) If $k = n$, then $\boldsymbol{A}$ is diagonalizable;
    (b) otherwise do next step.
  (2) If $s_1 = \cdots = s_l = 0$, then do next step; otherwise $\boldsymbol{A}$ is not diagonalizable.
  (3) For each eigenvalue λ_i whose $r_i > 1$, find the dimension of the eigenspace E_{λ_i}. If for each i, $r_i = \dim(E_{\lambda_i})$, then $\boldsymbol{A}$ is diagonalizable; otherwise $\boldsymbol{A}$ is not diagonalizable.

Exercise (6.6)

Let $\boldsymbol{A} = \begin{pmatrix} 0 & -1 & 0 \\ 2 & -3 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

(a) *Show that* -1 *is an eigenvalue of* $\boldsymbol{A}$.

(b) *Show that* $\dim(E_{-1}) = 2$.

(c) *Find a* 3×3 *matrix* $\boldsymbol{B}$ *such that* -3 *is an eigenvalue of* $\boldsymbol{BA}$.

Proof and Solution.

(a) Since $-\boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} -1 & 1 & \\ -2 & 2 & \\ & & 0 \end{pmatrix}$, we have $\det(-\boldsymbol{I} - \boldsymbol{A}) = 0$, and hence -1 is an eigenvalue of $\boldsymbol{A}$.

(b) Based on the Gaussian elimination, we will obtain the general solution for the linear system $(-\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ is $\boldsymbol{x} = s(1,1,0)^T + t(0,0,1)^T$. That is, $E_{-1} = \{s(1,1,0)^T + t(0,0,1)^T \mid s, t \in \mathbb{R}\} = \text{span}\{(1,1,0)^T, (0,0,1)^T\}$. Since $(1,1,0)^T$ and $(0,0,1)^T$ are linearly independent, they form a basis for E_{-1}. Hence $\dim(E_{-1}) = 2$.

(c) Take $\boldsymbol{B}$ to be $3\boldsymbol{I}_3$. Then $\det(-3\boldsymbol{I} - \boldsymbol{BA}) = \det(-3\boldsymbol{I} - 3\boldsymbol{IA}) = \det(-3\boldsymbol{I} - 3\boldsymbol{A}) = 3^3 \det(-\boldsymbol{I} - \boldsymbol{A}) = 0$, and hence -3 is an eigenvalue of $\boldsymbol{BA}$.

□

Exercise (6.10)

Let $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$.

(a) *Find a matrix* $\boldsymbol{P}$ *that diagonalizes* $\boldsymbol{A}$.

(b) *Compute* $\boldsymbol{A}^{10}$.

(c) *Find a matrix* $\boldsymbol{B}$ *such that* $\boldsymbol{B}^2 = \boldsymbol{A}$.

Solution of part (a).

Since $\boldsymbol{A}$ is an upper-triangular matrix, the all eigenvalues of $\boldsymbol{A}$ are $1, 4, 4$.

- For eigenvalue 1, the general solution for $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ is $s(1,0,0)^T$. So $E_1 = \{s(1,0,0)^T \mid s \in \mathbb{R}\}$, and we may take $\{(1,0,0)^T\}$ as a basis for E_1.
- For eigenvalue 4, the general solution for $(4\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ is $t(1,0,1)^T + v(0,1,0)^T$. So $E_4 = \{t(1,0,1)^T + v(0,1,0)^T \mid t, v \in \mathbb{R}\} = \operatorname{span}\{(1,0,1)^T, (0,1,0)^T\}$, and we may take $\{(1,0,1)^T, (0,1,0)^T\}$ as a basis for E_4.

Therefore $\boldsymbol{P} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ will diagonalize $\boldsymbol{A}$. $\square$

Solution of parts (b) and (c).

(b) (See Discussion 6.2.7) By part (a), we have $\boldsymbol{A} = \boldsymbol{P}\begin{pmatrix}1 & & \\ & 4 & \\ & & 4\end{pmatrix}\boldsymbol{P}^{-1}$. Hence we will have

$$\boldsymbol{A}^{10} = \boldsymbol{P}\begin{pmatrix}1 & & \\ & 4 & \\ & & 4\end{pmatrix}^{10}\boldsymbol{P}^{-1} = \boldsymbol{P}\begin{pmatrix}1^{10} & & \\ & 4^{10} & \\ & & 4^{10}\end{pmatrix}\boldsymbol{P}^{-1} = \begin{pmatrix}1 & 0 & 4^{10}-1 \\ 0 & 4^{10} & 0 \\ 0 & 0 & 4^{10}\end{pmatrix}.$$

(c) By part (a), we have

$$\boldsymbol{A} = \boldsymbol{P}\begin{pmatrix}1 & & \\ & 4 & \\ & & 4\end{pmatrix}\boldsymbol{P}^{-1} = \underbrace{\boldsymbol{P}\begin{pmatrix}1 & & \\ & 2 & \\ & & 2\end{pmatrix}\boldsymbol{P}^{-1}}_{\boldsymbol{B}}\underbrace{\boldsymbol{P}\begin{pmatrix}1 & & \\ & 2 & \\ & & 2\end{pmatrix}\boldsymbol{P}^{-1}}_{\boldsymbol{B}}.$$

So we may take $\boldsymbol{B}$ to be $\boldsymbol{P}\begin{pmatrix}1 & & \\ & 2 & \\ & & 2\end{pmatrix}\boldsymbol{P}^{-1} = \begin{pmatrix}1 & & 1 \\ & 2 & \\ & & 2\end{pmatrix}$, and $\boldsymbol{A} = \boldsymbol{B}^2$.

$\square$

Exercise (6.13(a))

Two square matrices A and B are said to be similar if there exists an invertible matrix P such that $P^{-1}AP = B$. Suppose A and B are similar matrices.

(i) *Show that A^n is similar to B^n for all positive integer n.*

(ii) *If A is invertible, show that B is invertible and A^{-1} is similar to B^{-1}.*

(iii) *If A is diagonalizable, show that B is diagonalizable.*

Proof of part (a).

(i) Since A and B are similar, then there exists an invertible matrix P, such that $P^{-1}AP = B$. Then for any positive integer n, we will have $P^{-1}A^nP = (P^{-1}AP)^n = B^n$, that is, A^n and B^n are similar.

(ii) Since $P^{-1}AP = B$, A and P are invertible, we have that B is invertible. And hence $B^{-1} = \mathbf{P^{-1}}AP^{-1} = P^{-1}A^{-1}P$, that is, A^{-1} and B^{-1} are similar.

(iii)
1. Since A is diagonalizable, there exists an invertible matrix Q, such that $Q^{-1}AQ$ is a diagonal matrix.
2. Since $A = PBP^{-1}$, we will have that $Q^{-1}PBP^{-1}Q$ is a diagonal matrix.
3. Let $R = P^{-1}Q$, then R is invertible. Therefore we will have that $R^{-1}BR$ is a diagonal matrix, that is, B is diagonalizable.

□

Exercise (6.13(b))

Show that $A = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & -1 \end{pmatrix}$ *and* $B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ *are similar.*

Proof.

1. Since A is an upper-triangular matrix, the all eigenvalues of A are $0, 1, -1$. Since this three eigenvalues are pairwise distinct, there exists an invertible matrix P, such that $P^{-1}AP = \begin{pmatrix} 0 & & \\ & 1 & \\ & & -1 \end{pmatrix}$.
2. By solving the equation $\det(\lambda I - A) = 0$, we get the all eigenvalues of B are $0, 1, -1$. Since this three eigenvalues are pairwise distinct, there exists an invertible matrix Q, such that $Q^{-1}BQ = \begin{pmatrix} 0 & & \\ & 1 & \\ & & -1 \end{pmatrix}$.
3. Let $R = PQ^{-1}$. Then $R^{-1}AR = QP^{-1}APQ^{-1} = Q\begin{pmatrix} 0 & & \\ & 1 & \\ & & -1 \end{pmatrix}Q^{-1} = B$, that is A and B are similar.

□

Exercise (6.14(a))

A square matrix $(a_{ij})_{n\times n}$ is called a stochastic matrix if all the entries are non-negative and the sum of entries of each column is 1, i.e. $a_{1i} + a_{2i} + \cdots + a_{ni} = 1$ for $i = 1, 2, \ldots, n$. Let $\boldsymbol{A}$ be a stochastic matrix.

(i) *Show that 1 is an eigenvalue of $\boldsymbol{A}$.*

(ii) *If λ is an eigenvalue of $\boldsymbol{A}$, then $|\lambda| \leq 1$.*

Proof of part (a-i).

$$\boldsymbol{A}^T \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} + a_{21} + \cdots + a_{n1} \\ a_{12} + a_{22} + \cdots + a_{n2} \\ \vdots \\ a_{1n} + a_{2n} + \cdots + a_{nn} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Thus 1 is an eigenvalue of $\boldsymbol{A}^T$. By Question 6.3, 1 is also an eigenvalue of $\boldsymbol{A}$. □

Proof of part (a-ii).

1. By Question 6.3, λ is an eigenvalue of $\boldsymbol{A}^T$.
2. Let $x = (x_1, x_2, \ldots, x_n)^T \neq \mathbf{0}$ be an eigenvector of $\boldsymbol{A}^T$ associated with the eigenvalue λ, that is, $\boldsymbol{A}^T\boldsymbol{x} = \lambda\boldsymbol{x}$.
3. Choose $k \in \{1, 2, \ldots, n\}$ such that $|x_k| = \max\limits_{i=1,2,\ldots,n} |x_i|$, that is, $|x_k| \geq |x_i|$ for $i = 1, 2, \ldots, n$. Since $\boldsymbol{x}$ is a non-zero vector, $|x_k| > 0$.
4. By comparing the k-th coordinate of both sides of $\boldsymbol{A}^T\boldsymbol{x} = \lambda\boldsymbol{x}$, we have
$$a_{1k}x_1 + a_{2k}x_2 + \cdots + a_{nk}x_n = \lambda x_k.$$
5. Hence we will have
$$\begin{aligned} |\lambda||x_k| &= |a_{1k}x_1 + a_{2k}x_2 + \cdots + a_{nk}x_n| \\ &\leq |a_{1k}x_1| + |a_{2k}x_2| + \cdots + |a_{nk}x_n| \\ &\leq a_{1k}|x_1| + a_{2k}|x_2| + \cdots + a_{nk}|x_n| && (a_{ij} \geq 0) \\ &\leq (a_{1k} + a_{2k} + \cdots + a_{nk})|x_k| = |x_k| \end{aligned}$$
Since $|x_k| > 0$, we have $|\lambda| \leq 1$.

$\square$

Exercise (6.14(b))

Let $B = \begin{pmatrix} 0.95 & 0 & 0 \\ 0.05 & 0.95 & 0.05 \\ 0 & 0.05 & 0.95 \end{pmatrix}$.

(i) *Is B a stochastic matrix?*

(ii) *Find a 3×3 invertible matrix P that diagonalizes B.*

Proof of part (b).

(i) All the entries are non-negative and the sum of entries of each column is 1, so B is a stochastic matrix.

(ii)
1. By solving the equation $\det(\lambda I - B) = 0$, the all eigenvalues of B are 1, 0.95 and 0.9.
2. It is easy to get $(0, 1, 1)^T$, $(-1, 0, 1)^T$ and $(0, -1, 1)^T$ are eigenvectors associated with 1, 0.95 and 0.9 respectively.
3. Let $P = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$, we will have $P^{-1}BP = \begin{pmatrix} 1 & & \\ & 0.95 & \\ & & 0.9 \end{pmatrix}$.

□

Exercise (6.16)

In a large city, the soft-drink market was 100% dominated by brand A. Four months ago, two new brands B and C were introduced to the market. According to the market research, for each month, about 1% and 2% of the customers of brand A switch to brands B and C respectively; and about 1% and 2% of the customers of brand B switch to brands A and C respectively; and about 2% and 2% of the customers of brand C switch to brands A and B respectively. Compute the present market shares of the three brands of soft-drink. Will the market shares stabilize in the long run if the trend continues? If so, estimate the market shares in the long run.

Solution.

(1) Let a_n, b_n and c_n be the percentage of customers choosing brand A, B and C, respectively, after n months. Then for any positive integer n,

$$\begin{cases} a_n = 0.97a_{n-1} + 0.01b_{n-1} + 0.02c_{n-1}; \\ b_n = 0.01a_{n-1} + 0.97b_{n-1} + 0.02c_{n-1}; \\ c_n = 0.02a_{n-1} + 0.02b_{n-1} + 0.96c_{n-1}. \end{cases}$$

□

Solution (Cont.)

(2) Let $\boldsymbol{x}_n = \begin{pmatrix} a_n \\ b_n \\ c_n \end{pmatrix}$ and $\boldsymbol{A} = \begin{pmatrix} 0.97 & 0.01 & 0.02 \\ 0.01 & 0.97 & 0.02 \\ 0.02 & 0.02 & 0.96 \end{pmatrix}$. Then the equations above can be represented by $\boldsymbol{x}_n = \boldsymbol{A}\boldsymbol{x}_{n-1} = \cdots = \boldsymbol{A}^n x_0$, where $\boldsymbol{x}_0 = \begin{pmatrix} 100 \\ 0 \\ 0 \end{pmatrix}$.

(3) By Algorithm 6.2.4, we find $\boldsymbol{P} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & 0 & 2 \end{pmatrix}$ such that
$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} 1 & & \\ & 0.96 & \\ & & 0.94 \end{pmatrix}.$$

(4) Then $\boldsymbol{x}_n = \boldsymbol{P}\begin{pmatrix} 1 & & \\ & 0.96^n & \\ & & 0.94^n \end{pmatrix}\boldsymbol{P}^{-1}\boldsymbol{x}_0 = \frac{50}{3}\begin{pmatrix} 2 + 3 \times 0.96^n + 0.94^n \\ 2 - 3 \times 0.96^n + 0.94^n \\ 2 - 2 \times 0.94^n \end{pmatrix}$.

(5) Therefore the present market shares are $\frac{50}{3}[2 + 3 \times 0.96^4 + 0.94^4]\% \simeq 88.8\%$, $\frac{50}{3}[2 - 3 \times 0.96^4 + 0.94^4]\% \simeq 3.9\%$ and $\frac{50}{3}[2 - 2 \times 0.94^4]\% \simeq 7.3\%$ for brand A, B and C, respectively.

□

Exercise (6.18)

Let d_n be the determinant of the following $n \times n$ matrix:

$$\begin{pmatrix} 3 & 1 & & & & & \\ 1 & 3 & 1 & & & 0 & \\ & 1 & 3 & \ddots & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & 3 & 1 & \\ & 0 & & & 1 & 3 & 1 \\ & & & & & 1 & 3 \end{pmatrix}$$

Show that $d_n = 3d_{n-1} - d_{n-2}$. Hence, or otherwise, find d_n.

Proof and Solution.

1. Use cofactor expansion along the first row:

$$d_n = 3\begin{vmatrix} 3 & 1 & & & 0 \\ 1 & 3 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 3 & 1 \\ & 0 & & 1 & 3 & 1 \\ & & & & 1 & 3 \end{vmatrix}_{(n-1)\times(n-1)} - \begin{vmatrix} 1 & 1 & & & 0 \\ & 3 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 3 & 1 \\ & 0 & & 1 & 3 & 1 \\ & & & & 1 & 3 \end{vmatrix}_{(n-1)\times(n-1)}$$

2. The first determinant above is d_{n-1}. By using cofactor expansion along the first column, we find that the second determinant is d_{n-2}. So $d_n = 3d_{n-1} - d_{n-2}$.

3. Note that $d_1 = 3$ and $d_2 = \begin{vmatrix} 3 & 1 \\ 1 & 3 \end{vmatrix} = 8$. By the procedure discussed in Example 6.2.9.2 or Example 6.2.12, we obtain

$$d_n = \frac{5+3\sqrt{5}}{10}\left(\frac{3+\sqrt{5}}{2}\right)^n - \frac{5-3\sqrt{5}}{10}\left(\frac{3-\sqrt{5}}{2}\right)^n.$$

□

Exercise (Remak 6.2.5.2)

Let λ_0 be an eigenvalue of matrix $\boldsymbol{A}$. Then the algebraic multiplicity of λ_0 is greater than or equal to the geometric multiplicity of λ_0.

Proof.

1. Assume $\dim(E_{\lambda_0}) = m$, then we can take a basis of E_{λ_0}: $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m\}$. Then we will get a basis for $\mathbb{R}^n$: $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m, \boldsymbol{\alpha}_{m+1}, \ldots, \boldsymbol{\alpha}_n\}$.
2. $$\boldsymbol{A}\begin{pmatrix}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_m & \cdots & \boldsymbol{\alpha}_n\end{pmatrix} = \begin{pmatrix}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_m & \cdots & \boldsymbol{\alpha}_n\end{pmatrix}\begin{pmatrix}\lambda_0 \boldsymbol{I}_m & \boldsymbol{B} \\ 0 & \boldsymbol{C}\end{pmatrix}.$$
3. Then $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - \lambda_0)^m \det(\lambda \boldsymbol{I}_{n-m} - \boldsymbol{C})$.
4. Hence, the algebraic multiplicity of some eigenvalue λ_0 is greater then or equal to the geometric multiplicity of λ_0.

$\square$

Exercise (Remark 6.2.5.3)

Let $\lambda_1, \lambda_2, \ldots, \lambda_t$, $t \geq 2$ be distinct eigenvalues of matrix $\boldsymbol{A}$, and $\boldsymbol{x}_i$ be the eigenvectors associated with λ_i, respectively. Then $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_t$ are linearly independent.

Proof: 1st Method.

1. First consider the case $t = 2$: if $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are linearly dependent, then there exist a, b, such that $a\boldsymbol{x}_1 + b\boldsymbol{x}_2 = \mathbf{0}$, where not both of a, b are zero.
2. Then $a\lambda_1\boldsymbol{x}_1 + b\lambda_2\boldsymbol{x}_2 = \boldsymbol{A}a\boldsymbol{x}_1 + \boldsymbol{A}b\boldsymbol{x}_2 = \boldsymbol{A}\mathbf{0} = \mathbf{0}$, and $a\lambda_1\boldsymbol{x}_1 + b\lambda_1\boldsymbol{x}_2 = \mathbf{0}$.
3. Then we will get $b(\lambda_1 - \lambda_2)\boldsymbol{x}_2 = \mathbf{0}$, i.e., $b = 0$. Similarly, $a = 0$. Contradiction.
4. For general case, we can apply mathematical induction, leave it for you.

□

Proof: 2nd Method.

(1) If $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_t$ are linearly dependent, then there exist some constant numbers $a_1, a_2, \ldots, a_t$, such that $a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_t\boldsymbol{x}_t = \mathbf{0}$, where not all of $a_1, \ldots, a_t$ are zero.

(2) Then $\mathbf{0} = \boldsymbol{A} \cdot \mathbf{0} = \boldsymbol{A}(a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_t\boldsymbol{x}_t) = a_1\lambda_1\boldsymbol{x}_1 + a_2\lambda_2\boldsymbol{x}_2 + \cdots + a_t\lambda_t\boldsymbol{x}_t$.

(3) Similarly, we have $a_1\lambda_1^2\boldsymbol{x}_1 + a_2\lambda_2^2\boldsymbol{x}_2 + \cdots + a_t\lambda_t^2\boldsymbol{x}_t = \mathbf{0}$.

(4) By induction, we have $a_1\lambda_1^j\boldsymbol{x}_1 + a_2\lambda_2^j\boldsymbol{x}_2 + \cdots + a_t\lambda_t^j\boldsymbol{x}_t = \mathbf{0}$ for $j = 1, 2, \ldots, t$.

(5) Consider the linear system: $\begin{cases} \lambda_1 y_1 + \lambda_2 y_2 + \cdots + \lambda_t y_t = \mathbf{0} \\ \lambda_1^2 y_1 + \lambda_2^2 y_2 + \cdots + \lambda_t^2 y_t = \mathbf{0} \\ \cdots \quad \cdots \\ \lambda_1^t y_1 + \lambda_2^t y_2 + \cdots + \lambda_t^t y_t = \mathbf{0} \end{cases}$

□

Proof: 2nd Method (Cont.)

(6) Let $\boldsymbol{x}_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{it} \end{pmatrix}$ for $i = 1, 2, \dots, t$.

(7) Then $(a_1 x_{1i}, a_2 x_{2i}, \dots, a_t x_{ti})^T$ satisfies that linear system, for all $i = 1, 2, \dots, t$.

(8) While $\det \begin{pmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_t \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_t^2 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^t & \lambda_2^t & \dots & \lambda_t^t \end{pmatrix} = \prod_{i=1}^{t} \lambda_i \prod_{1 \le i < j \le t} (\lambda_i - \lambda_j) \neq 0$. That is, that homogeneous linear system has only trivial zero solution.

(9) Since $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_t$ are nonzero vectors, $a_1 = a_2 = \cdots = a_t = 0$. Contradiction.

□

Exercise (6.3)

Let $\boldsymbol{A}$ be a square matrix and λ an eigenvalue of $\boldsymbol{A}$. Show that λ is an eigenvalue of $\boldsymbol{A}^T$.

Proof.

$$\begin{aligned}
&\lambda \text{ is an eigenvalue of } \boldsymbol{A} \\
\Leftrightarrow &\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0 \\
\Leftrightarrow &\det\left((\lambda \boldsymbol{I} - \boldsymbol{A})^T\right) = 0 \\
\Leftrightarrow &\det(\lambda \boldsymbol{I} - \boldsymbol{A}^T) = 0 \\
\Leftrightarrow &\lambda \text{ is an eigenvalue of } \boldsymbol{A}^T
\end{aligned}$$

□

Exercise (Question 6(b) in Final of 2006-2007(II))

If λ is an eigenvalue of a matrix $\boldsymbol{A}$, then $E_\lambda(\boldsymbol{A})$ and $E_\lambda(\boldsymbol{A}^T)$ of $\boldsymbol{A}$ and $\boldsymbol{A}^T$ have the same dimension.

Exercise (6.7(c))

Let $\boldsymbol{A} = \begin{pmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{pmatrix}$. *If* $\boldsymbol{B}$ *is another* 3×3 *matrix with an eigenvalue* λ *such that the dimension of the eigenspace associated with* λ *is 2, prove that* $2 + \lambda$ *is an eigenvalue of the matrix* $\boldsymbol{A} + \boldsymbol{B}$*.*

Proof of parts (a) and (b).

(a) Suppose $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - 2)^2(\lambda - 9) = 0$, then the eigenvalues are 2,2,9.

(b) Suppose $(2\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$, i.e. $\begin{pmatrix} -2 & 1 & -6 \\ -2 & 1 & -6 \\ -2 & 1 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \boldsymbol{0}$. A general solution is $t(1,2,0)^T + s(-3,0,1)^T$, i.e. $\{(1,2,0)^T, (-3,0,1)^T\}$ is a basis for the eigenspace associated with 2.

$\square$

Proof of part (c).

1. Let E_2 be the eigenspace of $\boldsymbol{A}$ associated with 2 and let E'_λ be the eigenspace of $\boldsymbol{B}$ associated with λ.
2. Since E_2 and E'_λ are subspaces of $\mathbb{R}^3$ and have dimension 2, they are two planes in $\mathbb{R}^3$ that contain the origin. So $E_2 \cap E'_\lambda$ is either a line through the origin or a plane containing the origin.
3. In both cases, we can find a nonzero vector $\boldsymbol{u} \in E_2 \cap E'_\lambda$, i.e. $\boldsymbol{Au} = 2\boldsymbol{u}$ and $\boldsymbol{Bu} = \lambda\boldsymbol{u}$, such that

$$(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{u} = \boldsymbol{Au} + \boldsymbol{Bu} = 2\boldsymbol{u} + \lambda\boldsymbol{u} = (2+\lambda)\boldsymbol{u}.$$

4. So $2+\lambda$ is an eigenvalue of $\boldsymbol{A}+\boldsymbol{B}$.

$\square$

Exercise (6.11)

Find a 3×3 matrix which has eigenvalue 1, 0, and -1 with corresponding eigenvectors $(0,1,1)^T$, $(1,-1,1)^T$ and $(1,0,0)^T$ respectively.

Proof.

Let $\boldsymbol{P} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$, and $\boldsymbol{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Then

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1} = \begin{pmatrix} -1 & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

satisfies the requirement. □

Remark

$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D}$	True
$\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1} = \boldsymbol{D}$	False

Exercise (6.12)

Determine the values of a and b so that the matrix $\begin{pmatrix} a & 1 \\ 0 & b \end{pmatrix}$ is diagonalizable.

Proof.

Claim: The matrix is diagonalizable if and only if $a \neq b$.

- If $a \neq b$, then there are 2 distinct eigenvalues, so the matrix is diagonalizable.
- If $a = b$, then consider the linear system $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{0}$. A general solution is $t(1,0)^T$, where t is a parameter. That is, the dimension of the eigenspace associated with a is 1. Hence, the matrix cannot be diagonalizable.

$\square$

Exercise (Question 5 in Final of 2004-2005(II))

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be 2 $n \times n$ diagonalizable matrices such that $\boldsymbol{AB} = \boldsymbol{BA}$. Prove that there exists an invertible matrix $\boldsymbol{P}$ such that $\boldsymbol{PAP}^{-1}$ and $\boldsymbol{PBP}^{-1}$ are both diagonal matrices.

Proof.

1. There exists an invertible $\boldsymbol{Q}$, such that, $\boldsymbol{C} = \boldsymbol{Q}^{-1}\boldsymbol{AQ} = \begin{pmatrix} \lambda_1 \boldsymbol{I}_{e_1} & & \\ & \ddots & \\ & & \lambda_t \boldsymbol{I}_{e_t} \end{pmatrix}$, where $\lambda_1, \dots, \lambda_t$ are all distinct eigenvalues of $\boldsymbol{A}$.
2. Let $\boldsymbol{D} = \boldsymbol{Q}^{-1}\boldsymbol{BQ}$, then $\boldsymbol{CD} = \boldsymbol{DC}$, and hence $\boldsymbol{D} = \begin{pmatrix} \boldsymbol{D}_{e_1} & & \\ & \ddots & \\ & & \boldsymbol{D}_{e_t} \end{pmatrix}$.
3. Since $\boldsymbol{B}$ is diagonalizable, so is $\boldsymbol{D}$, and hence so is $\boldsymbol{D}_{e_i}$ for all $i = 1, \dots, t$. Let $\boldsymbol{R}_i \boldsymbol{D}_{e_i} \boldsymbol{R}_i$ be a diagonal matrix.
4. Let $\boldsymbol{R} = \begin{pmatrix} \boldsymbol{R}_1 & & \\ & \ddots & \\ & & \boldsymbol{R}_t \end{pmatrix}$, then $\boldsymbol{R}^{-1}\boldsymbol{CR} = \begin{pmatrix} \lambda_1 \boldsymbol{I}_{e_1} & & \\ & \ddots & \\ & & \lambda_t \boldsymbol{I}_{e_t} \end{pmatrix}$ and $\boldsymbol{R}^{-1}\boldsymbol{DR}$ are diagonal.
5. Let $\boldsymbol{P} = \boldsymbol{QR}$, then $\boldsymbol{P}^{-1}\boldsymbol{AP}$ and $\boldsymbol{P}^{-1}\boldsymbol{BP}$ are diagonal.

$\square$

Exercise (Question 1(a) in Final of 2005-2006(I))

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

(i) *Write down the characteristic polynomial and eigenvalues of* $\boldsymbol{A}$.

(ii) *Write down the characteristic polynomial and eigenvalues of* $\boldsymbol{A}^5$.

(iii) *Is* $\boldsymbol{A}$ *diagonalizable?*

Exercise (Question 4(b) in Final of 2006-2007(II))

Let $\boldsymbol{B}$ *be a* 4×4 *matrix and* $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\}$ *a basis for* $\mathbb{R}^4$. *Suppose* $\boldsymbol{B}\boldsymbol{u}_1 = 2\boldsymbol{u}_1$, $\boldsymbol{B}\boldsymbol{u}_2 = \boldsymbol{0}$, $\boldsymbol{B}\boldsymbol{u}_3 = \boldsymbol{u}_4$, $\boldsymbol{B}\boldsymbol{u}_4 = \boldsymbol{u}_3$.

(i) *Find the eigenvalues of* $\boldsymbol{B}$.

(ii) *Find an eigenvalue that corresponds to each eigenvalue of* $\boldsymbol{B}$.

(iii) *Is* $\boldsymbol{B}$ *a diagonalizable matrix? Why?*

Hint.

$$\boldsymbol{B}\begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 & \boldsymbol{u}_4 \end{pmatrix} = \begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 & \boldsymbol{u}_4 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

$\square$

Exercise (Question 3(b-iii) in Final of 2009-2010(I))

For $n \geq 2$, let $\boldsymbol{B}_n = (b_{ij})$ be a square matrix of order n such that

$$b_{ij} = \begin{cases} 0, & i > j \text{ or } j > i+1; \\ 1, & j = i+1 \\ k, & i = j \end{cases}$$

where k is a real number. Prove that $\boldsymbol{B}_n$ is not diagonalizable for all $n \geq 2$.

Exercise

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be square matrices with order n. Then $\boldsymbol{AB}$ and $\boldsymbol{BA}$ have the same characteristic polynomial.

Proof.

1. There exist two invertible matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$, such that $\boldsymbol{A} = \boldsymbol{P}\begin{pmatrix} \boldsymbol{I}_r & \\ & \boldsymbol{0}_{n-r} \end{pmatrix}\boldsymbol{Q}$.
2. Let $\boldsymbol{QBP} = \begin{pmatrix} \boldsymbol{R}_1 & \boldsymbol{R}_2 \\ \boldsymbol{R}_3 & \boldsymbol{R}_4 \end{pmatrix}$, where $\boldsymbol{R}_1$ is an $r \times r$ matrix. Then

$$\begin{aligned}
\det(\lambda \boldsymbol{I} - \boldsymbol{AB}) &= \det(\lambda \boldsymbol{I} - \boldsymbol{P}\begin{pmatrix} \boldsymbol{I}_r & \\ & \boldsymbol{0}_{n-r} \end{pmatrix}\boldsymbol{QB}) \\
&= \det(\boldsymbol{P})\det(\lambda \boldsymbol{I} - \begin{pmatrix} \boldsymbol{I}_r & \\ & \boldsymbol{0}_{n-r} \end{pmatrix}\boldsymbol{QBP})\det(\boldsymbol{P}^{-1}) \\
&= \det(\lambda \boldsymbol{I} - \begin{pmatrix} \boldsymbol{R}_1 & \boldsymbol{R}_2 \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}) = \det(\lambda \boldsymbol{I}_r - \boldsymbol{R}_1)\det(\lambda \boldsymbol{I}_{n-r}) \\
\det(\lambda \boldsymbol{I} - \boldsymbol{BA}) &= \det(\lambda \boldsymbol{I} - \boldsymbol{BP}\begin{pmatrix} \boldsymbol{I}_r & \\ & \boldsymbol{0}_{n-r} \end{pmatrix}\boldsymbol{Q}) \\
&= \det(\boldsymbol{Q}^{-1})\det(\lambda \boldsymbol{I} - \boldsymbol{QBP}\begin{pmatrix} \boldsymbol{I}_r & \\ & \boldsymbol{0}_{n-r} \end{pmatrix})\det(\boldsymbol{Q}) \\
&= \det(\lambda \boldsymbol{I} - \begin{pmatrix} \boldsymbol{R}_1 & \boldsymbol{0} \\ \boldsymbol{R}_3 & \boldsymbol{0} \end{pmatrix}) = \det(\lambda \boldsymbol{I}_r - \boldsymbol{R}_1)\det(\lambda \boldsymbol{I}_{n-r})
\end{aligned}$$

Change log

- Page 190: Revise a typo: "$(-3)^3$" to "3^3";
- Page 191: Revise two typos: "$\boldsymbol{P} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$" to "$\boldsymbol{P} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$", and "$\boldsymbol{I} - \boldsymbol{A}$" to "$4\boldsymbol{I} - \boldsymbol{A}$".

Last modified: 19:14, March 25, 2011.

Schedule of Tutorial 9

- Any question about last tutorial
- Review concepts:
  - Inner product;
  - Orthogonal and orthonormal bases, Gram-Schmidt process;
  - Projection, least squares solution.
- Tutorial: 5.6, 5.8, 5.10, 5.12, 5.18, 5.19
- Additional material: 4.25(b), 5.9, Question 4 in Final of 2003–2004(II).

Inner Products

Let $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ be two vectors in $\mathbb{R}^n$.

- The inner product of $\boldsymbol{u}$ and $\boldsymbol{v}$: $\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$.
- The norm of $\boldsymbol{u}$: $\|\boldsymbol{u}\| = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$. Vectors of norm 1 are called unit vectors.
- The distance between $\boldsymbol{u}$ and $\boldsymbol{v}$ is $d(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\|$.
- The angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ is $\cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\|\,\|\boldsymbol{v}\|}\right)$.
- Relation between inner products and matrix products:
  - If $\boldsymbol{u}$ and $\boldsymbol{v}$ are written as row vectors, then $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}\boldsymbol{v}^T$.
  - If $\boldsymbol{u}$ and $\boldsymbol{v}$ are written as column vectors, then $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^T\boldsymbol{v}$.
- Let c be a scalar and $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ vectors in $\mathbb{R}^n$. Then
  1. $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$;
  2. $(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}$ and $\boldsymbol{w} \cdot (\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{w} \cdot \boldsymbol{u} + \boldsymbol{w} \cdot \boldsymbol{v}$;
  3. $(c\boldsymbol{u}) \cdot \boldsymbol{v} = \boldsymbol{u} \cdot (c\boldsymbol{v}) = c(\boldsymbol{u} \cdot \boldsymbol{v})$;
  4. $\|c\boldsymbol{u}\| = |c|\|\boldsymbol{u}\|$;
  5. $\boldsymbol{u} \cdot \boldsymbol{u} \geq 0$; and $\boldsymbol{u} \cdot \boldsymbol{u} = 0$ iff $\boldsymbol{u} = \boldsymbol{0}$.

Orthogonal and Orthonormal Bases

Def $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbb{R}^n$ are called orthogonal if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

Def A set S of vectors in $\mathbb{R}^n$ is called orthogonal if every pair of distinct vectors in S are orthogonal.

Def A set S of vectors in $\mathbb{R}^n$ is called orthonormal if S is orthogonal and every vector in S is a unit vector.

- If S is an orthogonal set of nonzero vectors in a vector space, then S is linearly independent. (By contrapositive)

Def
- A basis S for a vector space is called an orthogonal basis if S is orthogonal.
- A basis S for a vector space is called an orthonormal basis if S is orthonormal.

- Let V be a subspace of $\mathbb{R}^n$ and $\boldsymbol{w}$ a vector in V.
  - If $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$ is a basis for V, then

    $$\boldsymbol{w} = a_1\boldsymbol{w}_1 + a_2\boldsymbol{w}_2 + \cdots + a_k\boldsymbol{w}_k,$$

    where we need to solve linear system to get $a_1, a_2, \ldots, a_k$.
  - If $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ is an orthogonal basis for V, then

    $$\boldsymbol{w} = \frac{\boldsymbol{w} \cdot \boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\boldsymbol{u}_1 + \frac{\boldsymbol{w} \cdot \boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2}\boldsymbol{u}_2 + \cdots + \frac{\boldsymbol{w} \cdot \boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}\boldsymbol{u}_k.$$

  - If $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ is an orthonormal basis for V, then

    $$\boldsymbol{w} = (\boldsymbol{w} \cdot \boldsymbol{v}_1)\boldsymbol{v}_1 + (\boldsymbol{w} \cdot \boldsymbol{v}_2)\boldsymbol{v}_2 + \cdots + (\boldsymbol{w} \cdot \boldsymbol{v}_k)\boldsymbol{v}_k.$$

Orthogonal and Orthonormal Bases (Cont.)

Let V be a subspace of $\mathbb{R}^n$.

Def A vector $\boldsymbol{u} \in \mathbb{R}^n$ is said to be orthogonal to V if $\boldsymbol{u}$ is orthogonal to all vectors in V.

- Let V be a plane in $\mathbb{R}^3$ defined by the equation $ax + by + cz = 0$. Then $\boldsymbol{n} = (a, b, c)$ is orthogonal to V. The vector $\boldsymbol{n}$ is called a normal vector of V.
- If $V = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is a subspace of $\mathbb{R}^n$, then a vector $\boldsymbol{v} \in \mathbb{R}^n$ is orthogonal to V iff $\boldsymbol{v} \cdot \boldsymbol{u}_i = 0$ for $i = 1, 2, \ldots, k$.
- Let $V^\perp = \{\boldsymbol{u} \in \mathbb{R}^n \mid \boldsymbol{u} \text{ is orthogonal to } V\}$, then
  - $V^\perp$ is a subspace of $\mathbb{R}^n$;
  - $V \cap V^\perp = \{\boldsymbol{0}\}$;
  - $V + V^\perp = \mathbb{R}^n$;
  - if $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k, \ldots, \boldsymbol{u}_n\}$ is an orthogonal basis for $\mathbb{R}^n$, where $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is an orthogonal basis for V, then $\{\boldsymbol{u}_{k+1}, \boldsymbol{u}_{k+2}, \ldots, \boldsymbol{u}_n\}$ is an orthogonal basis for $V^\perp$;
  - $\dim(V) + \dim(V^\perp) = \dim(\mathbb{R}^n)$: let $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ be an orthogonal basis for $\boldsymbol{V}$, then it can be extended to an orthogonal basis for $\mathbb{R}^n$: $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k, \ldots, \boldsymbol{u}_n\}$. Then $\{\boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_n\}$ is an orthogonal basis for $V^\perp$.

Gram-Schmidt Process

Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ be a basis for a vector space V. Let

$$
\begin{aligned}
\boldsymbol{v}_1 &= \boldsymbol{u}_1 \\
\boldsymbol{v}_2 &= \boldsymbol{u}_2 - \frac{\boldsymbol{u}_2 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 \\
\boldsymbol{v}_3 &= \boldsymbol{u}_3 - \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 - \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_2}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 \\
&\vdots \\
\boldsymbol{v}_k &= \boldsymbol{u}_k - \frac{\boldsymbol{u}_k \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 - \frac{\boldsymbol{u}_k \cdot \boldsymbol{v}_2}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 - \cdots - \frac{\boldsymbol{u}_k \cdot \boldsymbol{v}_{k-1}}{\|\boldsymbol{v}_{k-1}\|^2}\boldsymbol{v}_{k-1}
\end{aligned}
$$

Then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ is an orthogonal basis for V. Furthermore, let $\boldsymbol{w}_i = \frac{\boldsymbol{v}_i}{\|\boldsymbol{v}_i\|}$ for $i = 1, 2, \ldots, k$. Then $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k\}$ is an orthonormal basis for V.

Projection

- Let V be a subspace of $\mathbb{R}^n$. Every vector $\boldsymbol{u} \in \mathbb{R}^n$ can be written uniquely as $\boldsymbol{u} = \boldsymbol{n} + \boldsymbol{p}$ such that $\boldsymbol{n}$ is a vector orthogonal to V and $\boldsymbol{p}$ is a vector in V. The vector $\boldsymbol{p}$ is called the (orthogonal) projection of $\boldsymbol{u}$ onto V
- Let V be a subspace of $\mathbb{R}^n$ and $\boldsymbol{w}$ a vector in $\mathbb{R}^n$.
  1. If $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ is an orthogonal basis for V, then
  $$\frac{\boldsymbol{w} \cdot \boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\boldsymbol{u}_1 + \frac{\boldsymbol{w} \cdot \boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2}\boldsymbol{u}_2 + \cdots + \frac{\boldsymbol{w} \cdot \boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}\boldsymbol{u}_k$$
  is the projection of $\boldsymbol{w}$ onto V.
  2. If $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ is an orthonormal basis for V, then
  $$(\boldsymbol{w} \cdot \boldsymbol{v}_1)\boldsymbol{v}_1 + (\boldsymbol{w} \cdot \boldsymbol{v}_2)\boldsymbol{v}_2 + \cdots + (\boldsymbol{w} \cdot \boldsymbol{v}_k)\boldsymbol{v}_k$$
  is the projection of $\boldsymbol{w}$ onto V.
- Let $\boldsymbol{p}$ be the projection of $\boldsymbol{u}$ onto V, then $\|\boldsymbol{u} - \boldsymbol{p}\| \leq \|\boldsymbol{u} - \boldsymbol{v}\|$ for any vector $\boldsymbol{v} \in V$, i.e. $\boldsymbol{p}$ is the best approximation of $\boldsymbol{u}$ in V.

Least Square Solution

- Let $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ be a linear system where $\boldsymbol{A}$ is an $m \times n$ matrix. A vector $\boldsymbol{x} \in \mathbb{R}^n$ is called the least squares solution to the linear system if it minimizes the value of $\|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|$.
- The following statements are equivalent:
  - $\boldsymbol{x}$ is the least squares solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$;
  - $\boldsymbol{x}$ is the solution $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{p}$ where $\boldsymbol{p}$ is the projection of $\boldsymbol{b}$ onto the column space of $\boldsymbol{A}$;
  - $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$.
- The linear system $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$ is always consistent:

$$\begin{aligned}
\operatorname{rank}(\boldsymbol{A}^T\boldsymbol{A} \mid \boldsymbol{A}^T\boldsymbol{b}) &= \operatorname{rank}(\boldsymbol{A}^T(\boldsymbol{A} \mid \boldsymbol{b})) \\
&\leq \min\{\operatorname{rank}(\boldsymbol{A}^T), \operatorname{rank}(\boldsymbol{A} \mid \boldsymbol{b})\} \\
&= \min\{\operatorname{rank}(\boldsymbol{A}), \operatorname{rank}(\boldsymbol{A} \mid \boldsymbol{b})\} \\
&= \operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}^T\boldsymbol{A})
\end{aligned}$$

- Suppose a linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is consistent. Then the solution set of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is equal to the solution set of $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$.

Exercise (5.6)

Let W be a subspace of $\mathbb{R}^n$. Define $W^\perp = \{\boldsymbol{u} \in \mathbb{R}^n \mid \boldsymbol{u}$ is orthogonal to $W\}$.

(a) *Let $W = \mathrm{span}\{(1,0,1,1),(1,-1,0,2),(1,2,3,-1)\}$. Find $W^\perp$.*

(b) *Show that $W^\perp$ is a subspace of $\mathbb{R}^n$.*

Solution of part (a).

1. Let (x, y, z, w) be any vector in $W^\perp$.
2. Then it is equivalent to

$$\begin{cases} (1,0,1,1)\cdot(x,y,z,w) = 0 \\ (1,-1,0,2)\cdot(x,y,z,w) = 0 \\ (1,2,3,-1)\cdot(x,y,z,w) = 0 \end{cases} \Leftrightarrow \begin{cases} x + z + w & = 0 \\ x - y + 2w & = 0 \\ x + 2y + 3z - w & = 0 \end{cases} \Leftrightarrow \begin{cases} x = -s - t \\ y = -s + t \\ z = s \\ w = t \end{cases}$$

   for some $s, t \in \mathbb{R}$.
3. So $W^\perp = \{s(-1,-1,1,0) + t(-1,1,0,1) \mid s, t \in \mathbb{R}\}$.

$\square$

Proof of part (b).

1. Let $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k\}$ be a basis for W.
2. Then we have

$$\boldsymbol{u} \in W^{\perp} \Leftrightarrow \begin{cases} \boldsymbol{w}_1 \cdot \boldsymbol{u} = 0 \\ \cdots \\ \boldsymbol{w}_k \cdot \boldsymbol{u} = 0 \end{cases} \Leftrightarrow \begin{pmatrix} \boldsymbol{w}_1 \\ \vdots \\ \boldsymbol{w}_k \end{pmatrix} \boldsymbol{u}^T = \boldsymbol{0}.$$

   Here we regard $\boldsymbol{u}$ as a row vector.
3. So $W^{\perp}$ is a solution set of a homogeneous system, and hence $W^{\perp}$ is a subspace of $\mathbb{R}^n$.

$\square$

Remark

We also may prove this by showing $W^{\perp}$ is non-empty and satisfies closed condition.

Exercise (5.8)

Let $ax + by + cz = d$ *be a plane in* $\mathbb{R}^3$*. Show that the vector* (a, b, c) *is perpendicular to the plane.*

Proof.

1. Note that $ax + by + cz = d$ is parallel to $ax + by + cz = 0$:
   - If $d = 0$, then they are same, and hence parallel;
   - If $d \neq 0$, then any point on the plane $ax + by + cz = d$ is not on the plane $ax + by + cz = 0$, vice versa. Since it is known that the relation between two planes in $\mathbb{R}^3$ has only 2 cases—intersection and parallelism, they are parallel.
2. Since (a, b, c) is perpendicular to $ax + by + cz = 0$, it would also be perpendicular to $ax + by + cz = d$.

□

Exercise (5.10)

For each of the following the line l and plane P in $\mathbb{R}^3$, determine whether l is perpendicular to P.

(a) *l: $x = 1 + 2t$, $y = t$, $z = 2 - t$ for $t \in \mathbb{R}$; P: $4x + 2y - 2z = 7$.*

(b) *l: $x = 1 + t$, $y = -1 + t$, $z = 3t$ for $t \in \mathbb{R}$; P: $2x + 2y = 5$.*

Solution.

(a) l can be represented as $(x, y, z) = (1, 0, 2) + t(2, 1, -1)$. So l is parallel to $(4, 2, -2) = 2(2, 1, -1)$. By Question 5.8, l is perpendicular to the plane $4x + 2y - 2z = 7$.

(b) l can be represented as $(x, y, z) = (1, -1, 0) + t(1, 1, 3)$. So l is parallel to $(1, 1, 3)$. On the other hand, a vector perpendicular to the plane $2x + 2y = 5$ must be parallel to $(2, 2, 0)$. Since $(2, 2, 0)$ and $(1, 1, 3)$ are not parallel, l is not perpendicular to the plane P.

$\square$

Exercise (5.12)

Let $\boldsymbol{u}_1 = (-2, -4, 1)$, $\boldsymbol{u}_2 = (3, -1, 2)$ *and* $\boldsymbol{u}_3 = (1, -1, -2)$.

(a) *Show that* $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ *is an orthogonal basis for* $\mathbb{R}^3$.

(b) *Let* $V = \operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ *and* $W = \operatorname{span}\{\boldsymbol{u}_3\}$. *Write each of the following vectors as a sum of two vectors* $\boldsymbol{v}$ *and* $\boldsymbol{w}$ *such that* $\boldsymbol{v} \in V$ *and* $\boldsymbol{w} \in W$: *(i)* $(0, 0, 1)$; *(ii)* $(1, 1, 0)$.

Proof and Solution.

(a) It is easy to check that $\boldsymbol{u}_i \cdot \boldsymbol{u}_j = 0$ for $i \neq j$.

(b) For any $\boldsymbol{x} \in \mathbb{R}^3$, by Theorem 5.2.8,

$$\boldsymbol{x} = \underbrace{\frac{\boldsymbol{x} \cdot \boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\boldsymbol{u}_1 + \frac{\boldsymbol{x} \cdot \boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2}\boldsymbol{u}_2}_{\boldsymbol{v}} + \underbrace{\frac{\boldsymbol{x} \cdot \boldsymbol{u}_3}{\|\boldsymbol{u}_3\|^2}\boldsymbol{u}_3}_{\boldsymbol{w}}.$$

Let $\boldsymbol{v} = \frac{\boldsymbol{x} \cdot \boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\boldsymbol{u}_1 + \frac{\boldsymbol{x} \cdot \boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2}\boldsymbol{u}_2 \in V$, and $\boldsymbol{w} = \frac{\boldsymbol{x} \cdot \boldsymbol{u}_3}{\|\boldsymbol{u}_3\|^2}\boldsymbol{u}_3 \in W$, then $\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{v}$. Hence following this process, we will have:

(i) $\boldsymbol{v} = (\frac{1}{3}, -\frac{1}{3}, \frac{1}{3})$ and $\boldsymbol{w} = (-\frac{1}{3}, \frac{1}{3}, \frac{2}{3})$.
(ii) $\boldsymbol{v} = (1, 1, 0)$ and $\boldsymbol{w} = (0, 0, 0)$.

□

Exercise (5.18)

Let $V = \text{span}\{(1,1,1),(1,p,p)\}$ *where* p *is a real number. Find an orthonormal basis for* V *and compute the projection of* $(5,3,1)$ *onto* V*.*

Solution.

- When $p=1$, $V = \text{span}\{(1,1,1)\}$ and hence $\left\{\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right)\right\}$ is an orthonormal basis for V. The projection of $(5,3,1)$ onto V is

$$\frac{(5,3,1)\cdot(1,1,1)}{\|(1,1,1)\|^2}(1,1,1) = (3,3,3).$$

- When $p \neq 1$. By observation, it is easy to obtain $V = \text{span}\{(1,1,1),(1,p,p)\} = \text{span}\{(1,0,0),(0,1,1)\}$. Since $(1,0,0)$ and $(0,1,1)$ are orthogonal, $\{(1,0,0),(0,1,1)\}$ is an orthogonal basis for V. Hence $\left\{(1,0,0),\left(0,\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)\right\}$ is an orthonormal basis for V. The projection of $(5,3,1)$ onto V is

$$\frac{(5,3,1)\cdot(1,0,0)}{\|(0,0,1)\|^2}(1,0,0) + \frac{(5,3,1)\cdot(0,1,1)}{\|(0,1,1)\|^2}(0,1,1) = (5,2,2).$$

$\square$

Exercise (5.19)

(a) *In $\mathbb{R}^2$, find the point on the line $y = x$ that is closet to the point $(1, 5)$.*

(b) *In $\mathbb{R}^2$, find the point on the line $y = x + 2$ that is closet to the point $(1, 5)$.*

Solution.

(a)
1. The line $y = x$ is a subspace in $\mathbb{R}^2$ spanned by $\{(1,1)\}$.
2. The projection of $(1, 5)$ onto the line $y = x$ is $\frac{(1,1)\cdot(1,5)}{\|(1,1)\|^2}(1, 1) = (3, 3)$.
3. By Theorem 5.3.2, we have that $(3, 3)$ is the point on $y = x$ that is closest to $(1, 5)$.

(b) Since $y - 2 = x$ is not a subspace, we cannot apply the method in part (a) directly.
1. If we move the line $y - 2 = x$ and the point $(1, 5)$ down by 2 in the y direction, the resultants are the line $y = x$ and the point $(1, 3)$.
2. By the method in part (a), $(2, 2)$ is the point on $y = x$ that is closest to $(1, 3)$.
3. Moving back, we obtain that $(2, 4)$ is the point on $y - 2 = x$ that is closest to $(1, 5)$.

□

Exercise (4.25(b))

Suppose a linear system $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent. Show that the solution set of $\boldsymbol{Ax} = \boldsymbol{b}$ is equal to the solution set of $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$.

Proof.

1. Let $\boldsymbol{v}$ be a solution of $\boldsymbol{Ax} = \boldsymbol{b}$.
2. Since $\boldsymbol{A}^T\boldsymbol{Av} = \boldsymbol{A}^T\boldsymbol{b}$, $\boldsymbol{v}$ is also a solution of $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$.
3. Since the nullspace of $\boldsymbol{A}$ and the nullspace of $\boldsymbol{A}^T\boldsymbol{A}$ are identical, we have

$$\begin{aligned}
\text{The solution set of } (\boldsymbol{Ax} = \boldsymbol{b}) &= \{\boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u} \in \text{nullspace of } (\boldsymbol{A})\} \\
&= \{\boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u} \in \text{nullspace of } (\boldsymbol{A}^T\boldsymbol{A})\} \\
&= \text{The solution set of } (\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b})
\end{aligned}$$

□

Exercise (Remark)

Uniqueness of the decomposition $\boldsymbol{u} = \boldsymbol{n} + \boldsymbol{p}$, where $\boldsymbol{n}$ is orthogonal the subspace V, and $\boldsymbol{p} \in V$.

Proof.

Proof by contradiction:

1. Assume that the decomposition is not unique. Then there exist $\boldsymbol{n}_1$, $\boldsymbol{n}_2$, $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$, such that $\boldsymbol{n}_1 + \boldsymbol{p}_1 = \boldsymbol{u} = \boldsymbol{n}_2 + \boldsymbol{p}_2$, where $\boldsymbol{n}_1, \boldsymbol{n}_2$ are orthogonal to V, and $\boldsymbol{p}_1, \boldsymbol{p}_2 \in V$.
2. Then we have
$$\boldsymbol{n}_1 - \boldsymbol{n}_2 = \boldsymbol{p}_2 - \boldsymbol{p}_1.$$
3. Since $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are orthogonal to V, so is $\boldsymbol{u}_1 - \boldsymbol{u}_2$, and hence $\boldsymbol{p}_2 - \boldsymbol{p}_1$ is also orthogonal to V.
4. Since $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are in the subspace V, we have $\boldsymbol{p}_2 - \boldsymbol{p}_1 \in V$.
5. Therefore $\boldsymbol{p}_2 - \boldsymbol{p}_1$ is orthogonal to itself, that is, $(\boldsymbol{p}_2 - \boldsymbol{p}_1) \cdot (\boldsymbol{p}_2 - \boldsymbol{p}_1) = 0$. Hence $\boldsymbol{p}_1 = \boldsymbol{p}_2$, and $\boldsymbol{n}_1 = \boldsymbol{n}_2$. Contradiction.

□

Exercise (5.9)

Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n\}$ *be an orthogonal set of vectors in a vector space. Show that*

$$\|\boldsymbol{u}_1 + \boldsymbol{u}_2 + \cdots + \boldsymbol{u}_n\|^2 = \|\boldsymbol{u}_1\|^2 + \|\boldsymbol{u}_2\|^2 + \cdots + \|\boldsymbol{u}_n\|^2.$$

For $n = 2$*, interpret the result geometrically in* $\mathbb{R}^2$*.*

Proof.

$$\begin{aligned}\|\boldsymbol{u}_1 + \boldsymbol{u}_2 + \cdots + \boldsymbol{u}_n\|^2 &= (\boldsymbol{u}_1 + \boldsymbol{u}_2 + \cdots + \boldsymbol{u}_n) \cdot (\boldsymbol{u}_1 + \boldsymbol{u}_2 + \cdots + \boldsymbol{u}_n) \\ &= (\boldsymbol{u}_1 \cdot \boldsymbol{u}_1) + \cdots + (\boldsymbol{u}_n \cdot \boldsymbol{u}_n) \qquad \text{Since } \boldsymbol{u}_i \cdot \boldsymbol{u}_j = 0 \text{ for } i \neq j \\ &= \|\boldsymbol{u}_1\|^2 + \|\boldsymbol{u}_2\|^2 + \cdots + \|\boldsymbol{u}_n\|^2\end{aligned}$$

For $n = 2$, it is Pythagoras' Theorem or 勾股定理. □

Exercise (Question 4 in Final of 2003-2004(II))

Let W be a subspace of $\mathbb{R}^n$ and let $W^\perp = \{\boldsymbol{u} \in \mathbb{R}^n \mid \boldsymbol{u}$ is orthogonal to $W\}$. Then

(i) *$W^\perp$ is a subspace of $\mathbb{R}^n$;*

(ii) $\dim(W) + \dim(W^\perp) = n$.

Change log

- Page 219: Revise typos: "$\dim(V) \cap \dim(V^{\perp})$" to "$V \cap V^{\perp}$", "$\dim(V) + \dim(V^{\perp})$" to "$V + V^{\perp}$";
- Page 219: Add a proof for "$\dim(V) + \dim(V^{\perp}) = n = \dim(\mathbb{R}^n)$";
- Page 220: Revise a typo: "$\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_k\}$ is an orthogonal basis" to "$\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_k\}$ is an orthonormal basis";
- Page 224: Add a remark "We also may prove this by showing $W^{\perp}$ is non-empty and satisfies closed condition";
- Page 228: Revise the proof;
- Page 231: Add a proof for the Remark.

Last modified: 22:00, April 3, 2011.

Schedule of Tutorial 10

- Any question about last tutorial
- Review concepts
  - Orthogonal matrices;
  - Symmetric matrices.
- Tutorial: 5.25, 5.29, 5.30, 5.32, 6.21, 6.22
- Additional material:
  - 2 additional equivalent statements for orthogonal matrices;
  - Problem 6.3.8;
  - Any eigenvalue of a symmetric matrix is a real number;
  - Exercise 5.33;
  - Question 5(3-6) in Final 2005–2006(I);
  - Question 6 in Final 2001–2002(II).

Orthogonal Matrices and Symmetric Matrices

- A square matrix $\boldsymbol{A}$ is called orthogonal if $\boldsymbol{A}^{-1} = \boldsymbol{A}^T$.
- $\boldsymbol{A}$ is a square matrix, then the following statements are equivalent:
  - $\boldsymbol{A}$ is orthogonal;
  - $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I}$;
  - $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}$;
  - the rows of $\boldsymbol{A}$ form an orthonormal basis for $\mathbb{R}^n$;
  - the columns of $\boldsymbol{A}$ form an orthonormal basis for $\mathbb{R}^n$;
  - $\|\boldsymbol{A}\boldsymbol{x}\| = \|\boldsymbol{x}\|$ for any vector $\boldsymbol{x} \in \mathbb{R}^n$;
  - $\boldsymbol{A}\boldsymbol{u} \cdot \boldsymbol{A}\boldsymbol{v} = \boldsymbol{u} \cdot \boldsymbol{v}$ for any vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$.
- Let $\boldsymbol{A}$ be an orthogonal matrix, λ an eigenvalue of $\boldsymbol{A}$, then $|\lambda| = 1$: Since $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ and $\|\boldsymbol{A}\boldsymbol{x}\| = \|\boldsymbol{x}\|$, we have $|\lambda| = 1$.
- Let $\boldsymbol{A}$ be a symmetric matrix. If $\boldsymbol{u}$ and $\boldsymbol{v}$ are two eigenvectors of $\boldsymbol{A}$ associated with eigenvalues λ and μ, respectively, where $\lambda \neq \mu$, show that $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.
- Let $\boldsymbol{A}$ be a symmetric matrix, λ an eigenvalue of $\boldsymbol{A}$, then λ is a real number.

Orthogonal diagonalization

Let $\boldsymbol{A}$ be a real matrix.

- If $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^T$, then $\boldsymbol{A}$ is called normal matrix.
- $\boldsymbol{A}$ is called orthogonally diagonalizable if there exists an orthogonal matrix $\boldsymbol{P}$ (real matrix) such that $\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}$ is a diagonal matrix.
- Let $\boldsymbol{A}$ be a normal matrix, and $a_1 \pm \sqrt{-1}b_1, \ldots, a_t \pm \sqrt{-1}b_t, \lambda_{2t+1}, \ldots, \lambda_n$ be all eigenvalues of $\boldsymbol{A}$, where $b_1, \ldots, b_t > 0$. Then $\boldsymbol{A}$ is orthogonally similar with

$$\boldsymbol{B} = \operatorname{diag}\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \cdots, \begin{pmatrix} a_t & b_t \\ -b_t & a_t \end{pmatrix}, \lambda_{2t+1}, \cdots, \lambda_n \right).$$

- If $\boldsymbol{A}$ is an orthogonal matrix, then $\boldsymbol{A}$ is orthogonally similar with

$$\boldsymbol{B} = \operatorname{diag}\left(\begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{pmatrix}, \cdots, \begin{pmatrix} \cos\theta_t & \sin\theta_t \\ -\sin\theta_t & \cos\theta_t \end{pmatrix}, \boldsymbol{I}_u, -\boldsymbol{I}_v \right),$$

  where $2t + u + v = n$, $0 < \theta_1 \leq \cdots \leq \theta_t < \pi$.
- If $\boldsymbol{A}$ is a symmetric matrix, then $\boldsymbol{A}$ is orthogonally similar with $\operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are all eigenvalues of $\boldsymbol{A}$. Furthermore, every symmetric matrix has n real eigenvalues.

Exercise (5.25(a))

(a) *Let* $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 0 \end{pmatrix}$ *and* $\boldsymbol{b} = \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}$.

- (i) *Solve the linear system* $\boldsymbol{Ax} = \boldsymbol{b}$.
- (ii) *Find the least squares solution to* $\boldsymbol{Ax} = \boldsymbol{b}$.

Solution of part (a).

(i) By observation, $x_1 = 2$, and then $x_2 = 1$. Hence $\boldsymbol{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

(ii) As we known, $\boldsymbol{x}$ is a least squares solution to $\boldsymbol{Ax} = \boldsymbol{b}$ iff $\boldsymbol{x}$ is a solution to $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$, so we only need to solve the following linear system

$$\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}.$$

By Gaussian elimination, we have the solution is $\boldsymbol{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, which is exact the least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$.

□

Exercise (5.25(b))

(b) *Suppose a linear system $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent. Show that the solution set of $\boldsymbol{Ax} = \boldsymbol{b}$ is equal to the solution set of $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$.*

Recall

- Theorem 4.3.5: If $\boldsymbol{v}$ is a solution of $\boldsymbol{Ax} = \boldsymbol{b}$, then the solution of $\boldsymbol{Ax} = \boldsymbol{b}$ is

$$\{\boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u} \in \text{the nullspace of } \boldsymbol{A}\}.$$

- Question 4.25(a): The nullspace of $\boldsymbol{A}$ is equal to the nullspace of $\boldsymbol{A}^T\boldsymbol{A}$.

Proof of part (b).

1. Let $\boldsymbol{v}$ be a solution of $\boldsymbol{Ax} = \boldsymbol{b}$.
2. Since $\boldsymbol{A}^T\boldsymbol{Av} = \boldsymbol{A}^T\boldsymbol{b}$, $\boldsymbol{v}$ is also a solution of $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$.
3. Since the nullspace of $\boldsymbol{A}$ and the nullspace of $\boldsymbol{A}^T\boldsymbol{A}$ are identical, we have

$$\begin{aligned}\text{The solution set of } (\boldsymbol{Ax} = \boldsymbol{b}) &= \{\boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u} \in \text{nullspace of } (\boldsymbol{A})\} \\ &= \{\boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u} \in \text{nullspace of } (\boldsymbol{A}^T\boldsymbol{A})\} \\ &= \text{The solution set of } (\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b})\end{aligned}$$

□

Exercise (5.29(a))

(a) *Let* $S_1 = \{(1,0),(0,1)\}$, $S_2 = \{(1,-1),(2,1)\}$ *and* $S_3 = \left\{(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})\right\}$. *Clearly,* S_1, S_2 *and* S_3 *are three bases for* $\mathbb{R}^2$*. Let* $\boldsymbol{u} = (1,4)$ *and* $\boldsymbol{v} = (-1,1)$*. Compute* $(\boldsymbol{u})_{S_i}$, $(\boldsymbol{v})_{S_i}$ *and* $(\boldsymbol{u})_{S_i} \cdot (\boldsymbol{v})_{S_i}$ *for* $i = 1,2,3$*. What do you observe?*

Solution of part (a).

- Since $S_1 = \{(1,0),(0,1)\}$ is the standard basis for $\mathbb{R}^2$, we have $(\boldsymbol{u})_{S_1} = \boldsymbol{u} = (1,4)$, $(\boldsymbol{v})_{S_1} = \boldsymbol{v} = (-1,1)$, and $(\boldsymbol{u})_{S_1} \cdot (\boldsymbol{v})_{S_1} = (1,4)\cdot(-1,1) = 3$.
- It is clear that S_3 is an orthonormal basis for $\mathbb{R}^2$. Thus $(\boldsymbol{u})_{S_3} = \left(\boldsymbol{u}\cdot(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), \boldsymbol{u}\cdot(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})\right) = \left(\frac{5}{\sqrt{2}}, \frac{3}{\sqrt{2}}\right)$,
  $(\boldsymbol{v})_{S_3} = \left(\boldsymbol{v}\cdot(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), \boldsymbol{v}\cdot(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})\right) = (0, \sqrt{2})$, and $(\boldsymbol{u})_{S_3} \cdot (\boldsymbol{v})_{S_3} = 3$.
- Since $(1,-1)$ and $(2,1)$ are not orthogonal, we can not use inner product to get the coordinate vectors. Assume $\boldsymbol{u} = a_1(1,-1) + a_2(2,1)$, by solving this linear system, we have $a_1 = -\frac{7}{3}$, $a_2 = \frac{5}{3}$, and hence $(\boldsymbol{u})_{S_2} = (-\frac{7}{3}, \frac{5}{3})$. Similarly, we will have $(\boldsymbol{v})_{S_2} = (-1,0)$. Hence $(\boldsymbol{u})_{S_2} \cdot (\boldsymbol{v})_{S_2} = \frac{7}{3}$.

Note that $(\boldsymbol{u})_{S_1} \cdot (\boldsymbol{v})_{S_1} = (\boldsymbol{u})_{S_3} \cdot (\boldsymbol{v})_{S_3} \neq (\boldsymbol{u})_{S_2} \cdot (\boldsymbol{v})_{S_2}$. See part (b) for an explanation. □

Exercise (5.29(b))

(b) *Prove that if S and T are two othornormal bases for a vector space V, then for any vectors $\boldsymbol{u}, \boldsymbol{v} \in V$, $(\boldsymbol{u})_S \cdot (\boldsymbol{v})_S = (\boldsymbol{u})_T \cdot (\boldsymbol{v})_T$.*

Proof of part (b).

1. Let $\boldsymbol{P}$ be the transition matrix from S to T. Since S and T are orthonormal bases, $\boldsymbol{P}$ is orthogonal, i.e. $\boldsymbol{P}^T\boldsymbol{P} = \boldsymbol{I}$:

$$(\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n) = (\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n)\boldsymbol{P}, \quad \text{see page 100 in textbook.}$$

2. By definition of the inner product, we have $[\boldsymbol{u}]_S \cdot [\boldsymbol{v}]_S = (\boldsymbol{u})_S \cdot (\boldsymbol{v})_S$.
3. Then we have

$$\begin{aligned}[\boldsymbol{u}]_T \cdot [\boldsymbol{v}]_T = ([\boldsymbol{u}]_T)^T[\boldsymbol{v}]_T = (\boldsymbol{P}[\boldsymbol{u}]_S)^T(\boldsymbol{P}[\boldsymbol{v}]_S) &= ([\boldsymbol{u}]_S)^T\boldsymbol{P}^T\boldsymbol{P}[\boldsymbol{v}]_S \\ &= ([\boldsymbol{u}]_S)^T[\boldsymbol{v}]_S = [\boldsymbol{u}]_S \cdot [\boldsymbol{v}]_S\end{aligned}$$

Therefore, we have $(\boldsymbol{u})_S \cdot (\boldsymbol{v})_S = (\boldsymbol{u})_T \cdot (\boldsymbol{v})_T$. $\square$

Exercise (5.30)

Let $\boldsymbol{A}$ be an orthogonal matrix of order n and let $\boldsymbol{u}$, $\boldsymbol{v}$ be any two vectors in $\mathbb{R}^n$. Show that

(a) $\|\boldsymbol{u}\| = \|\boldsymbol{A}\boldsymbol{u}\|$;

(b) $d(\boldsymbol{u}, \boldsymbol{v}) = d(\boldsymbol{A}\boldsymbol{u}, \boldsymbol{A}\boldsymbol{v})$;

(c) *the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ is equal to the angle between $\boldsymbol{A}\boldsymbol{u}$ and $\boldsymbol{A}\boldsymbol{v}$.*

Proof.

(a) $\|\boldsymbol{A}\boldsymbol{u}\|^2 = (\boldsymbol{A}\boldsymbol{u})^T(\boldsymbol{A}\boldsymbol{u}) = \boldsymbol{u}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{u} = \boldsymbol{u}^T\boldsymbol{u} = \|\boldsymbol{u}\|^2$. Since both $\|\boldsymbol{u}\|$ and $\|\boldsymbol{A}\boldsymbol{u}\|$ are nonnegative, we have $\|\boldsymbol{A}\boldsymbol{u}\| = \|\boldsymbol{u}\|$.

(b) By part (a), $d(\boldsymbol{A}\boldsymbol{u}, \boldsymbol{A}\boldsymbol{v}) = \|\boldsymbol{A}\boldsymbol{u} - \boldsymbol{A}\boldsymbol{v}\| = \|\boldsymbol{A}(\boldsymbol{u} - \boldsymbol{v})\| = \|\boldsymbol{u} - \boldsymbol{v}\| = d(\boldsymbol{u}, \boldsymbol{v})$.

(c) $(\boldsymbol{A}\boldsymbol{u}) \cdot (\boldsymbol{A}\boldsymbol{v}) = (\boldsymbol{A}\boldsymbol{u})^T\boldsymbol{A}\boldsymbol{v} = \boldsymbol{u}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{v} = \boldsymbol{u}^T\boldsymbol{v} = \boldsymbol{u} \cdot \boldsymbol{v}$. So the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ is

$$\cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right) = \cos^{-1}\left(\frac{\boldsymbol{A}\boldsymbol{u} \cdot \boldsymbol{A}\boldsymbol{v}}{\|\boldsymbol{A}\boldsymbol{u}\|\|\boldsymbol{A}\boldsymbol{v}\|}\right),$$

which is the angle between $\boldsymbol{A}\boldsymbol{u}$ and $\boldsymbol{A}\boldsymbol{v}$.

□

Exercise (5.32)

Let $\boldsymbol{A}$ be an orthogonal matrix of order n and let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ be a basis for $\mathbb{R}^n$.

(a) *Show that $T = \{\boldsymbol{A}\boldsymbol{u}_1, \boldsymbol{A}\boldsymbol{u}_2, \ldots, \boldsymbol{A}\boldsymbol{u}_n\}$ is a basis for $\mathbb{R}^n$.*

(b) *If S is orthogonal, show that T is orthogonal.*

(c) *If S is orthonormal, is T orthonormal?*

Proof and Solution.

(a) Since $\boldsymbol{A}$ is invertible, by Question 3.23(b)(i), T is linearly independent. So T is a basis for $\mathbb{R}^n$ by Theorem 3.5.6.

(b) By Question 5.30, for $i \neq j$ we know that the angle between $\boldsymbol{A}\boldsymbol{u}_i$ and $\boldsymbol{A}\boldsymbol{u}_j$ is same as the angle between $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$ which is $90°$, hence $\boldsymbol{A}\boldsymbol{u}_i$ and $\boldsymbol{A}\boldsymbol{u}_j$ are orthogonal, so T is orthogonal.

(c) Yes. By part (b), we know that T is orthogonal. Also by Question 5.30, since $\|\boldsymbol{A}\boldsymbol{u}_i\| = \|\boldsymbol{u}_i\| = 1$ for any $i = 1, 2, \ldots, n$, we have T is orthonormal.

□

Exercise (6.21)

Let $\boldsymbol{u}$ be a column matrix.

(a) *Show that $\boldsymbol{I}-\boldsymbol{u}\boldsymbol{u}^T$ is orthogonally diagonalizable.*

(b) *Find a matrix $\boldsymbol{P}$ that orthogonally diagonalizes $\boldsymbol{I}-\boldsymbol{u}\boldsymbol{u}^T$ if $\boldsymbol{u}=(1,-1,1)^T$.*

Proof.

(a) Since $\left(\boldsymbol{I}-\boldsymbol{u}\boldsymbol{u}^T\right)^T=\boldsymbol{I}-\boldsymbol{u}\boldsymbol{u}^T$, $\boldsymbol{I}-\boldsymbol{u}\boldsymbol{u}^T$ is symmetric. Hence $\boldsymbol{I}-\boldsymbol{u}\boldsymbol{u}^T$ is orthogonally diagonalizable.

(b) When $\boldsymbol{u}=(1,-1,1)^T$, $\boldsymbol{I}-\boldsymbol{u}\boldsymbol{u}^T=\begin{pmatrix}0&1&-1\\1&0&1\\-1&1&0\end{pmatrix}$. We will obtain all the eigenvalues are $1,1,-2$.

- For eigenvalue 1, by solving the linear system $(\boldsymbol{I}-[\boldsymbol{I}-\boldsymbol{u}\boldsymbol{u}^T])\boldsymbol{x}=\boldsymbol{0}$, we will have the general solution $\boldsymbol{x}=s(1,1,0)^T+t(-1,0,1)^T$. Hence, for the eigenspace E_1, we may take $\{(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0),(-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}})\}$ as an orthonormal basis.
- For eigenvalue -2, by solving the linear system $(-2\boldsymbol{I}-[\boldsymbol{I}-\boldsymbol{u}\boldsymbol{u}^T])\boldsymbol{x}=\boldsymbol{0}$, we will have the general solution $\boldsymbol{x}=s(1,-1,1)^T$. Hence, for the eigenspace E_{-2}, we may take $\{(\frac{1}{\sqrt{3}},-\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}})\}$ as an orthonormal basis.

Thus take $\boldsymbol{P}=\begin{pmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{6}}&-\frac{1}{\sqrt{3}}\\0&\frac{2}{\sqrt{6}}&\frac{1}{\sqrt{3}}\end{pmatrix}$, then $\boldsymbol{P}^T[\boldsymbol{I}-\boldsymbol{u}\boldsymbol{u}^T]\boldsymbol{P}=\begin{pmatrix}1&&\\&1&\\&&-2\end{pmatrix}$.

□

Exercise (6.22(ab))

Let $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$*, and* $\boldsymbol{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

(a) *Show that* $\boldsymbol{u}$ *is an eigenvector of* $\boldsymbol{A}$*.*

(b) *Let* $\boldsymbol{v} = (a, b, c, d)^T$*. Show that if* $\boldsymbol{u} \cdot \boldsymbol{v} = 0$*, then* $\boldsymbol{v}$ *is an eigenvector of* $\boldsymbol{A}$*.*

Proof of parts (a) and (b).

(a) Since $\boldsymbol{Au} = \begin{pmatrix} 4 \\ 4 \\ 4 \\ 4 \end{pmatrix} = 4\boldsymbol{u}$, $\boldsymbol{u}$ is an eigenvector associated with eigenvalue 4.

(b) Since $\boldsymbol{u} \cdot \boldsymbol{v} = 0$, we have $a + b + c + d = 0$. Therefore $\boldsymbol{Av} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0\boldsymbol{v}$, that is, $\boldsymbol{v}$ is an eigenvector associated with eigenvalue 0.

□

Exercise (6.22(c))

Let $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$, *and* $\boldsymbol{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

(c) *Suppose* $\boldsymbol{P} = \begin{pmatrix} \frac{1}{2} & a_1 & a_2 & a_3 \\ \frac{1}{2} & b_1 & b_2 & b_3 \\ \frac{1}{2} & c_1 & c_2 & c_3 \\ \frac{1}{2} & d_1 & d_2 & d_3 \end{pmatrix}$ *is an orthogonal matrix. Find* $\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}$.

Proof of part (c).

1. Since $\boldsymbol{P}$ is an orthogonal matrix, the columns form an orthonormal basis for $\mathbb{R}^4$. Thus $a_i + b_i + c_i + d_i = 0$ for $i = 1, 2, 3$.
2. By part (a), the first column of $\boldsymbol{P}$ is the eigenvector of $\boldsymbol{A}$ associated with the eigenvalue 4. By part (b), the other three columns of $\boldsymbol{P}$ are eigenvectors of $\boldsymbol{A}$ associated with the eigenvalue 0. So

$$\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

□

Exercise

Let $\boldsymbol{A}$ be a square matrix of order n. Then the following statements are equivalent:

(a) *$\boldsymbol{A}$ is an orthogonal matrix;*

(b) *$\|\boldsymbol{x}\| = \|\boldsymbol{A}\boldsymbol{x}\|$ for any vector $\boldsymbol{x} \in \mathbb{R}^n$;*

(c) *$\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{A}\boldsymbol{u} \cdot \boldsymbol{A}\boldsymbol{v}$ for any vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$.*

Proof.

- "(a)⇒(b)": $\|\boldsymbol{A}\boldsymbol{x}\|^2 = (\boldsymbol{A}\boldsymbol{x})^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{x} = \|\boldsymbol{x}\|^2$.
- "(b)⇒(c)": Since $\|\boldsymbol{u}+\boldsymbol{v}\| = \|\boldsymbol{A}(\boldsymbol{u}+\boldsymbol{v})\|$ and $\|\boldsymbol{u}-\boldsymbol{v}\| = \|\boldsymbol{A}(\boldsymbol{u}-\boldsymbol{v})\|$, we will get
$$\boldsymbol{u}^T\boldsymbol{v} + \boldsymbol{v}^T\boldsymbol{u} = \boldsymbol{u}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{v} + \boldsymbol{v}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{u}.$$
Since $\boldsymbol{u}^T\boldsymbol{v} = \boldsymbol{v}^T\boldsymbol{u}$ and $\boldsymbol{u}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{v} = \boldsymbol{v}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{u}$, we have
$$\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^T\boldsymbol{v} = \boldsymbol{u}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{v} = \boldsymbol{A}\boldsymbol{u} \cdot \boldsymbol{A}\boldsymbol{v}.$$
- "(c)⇒(a)": Take $\boldsymbol{u} = e_i$, $\boldsymbol{v} = e_j$, left is easy.

□

Exercise (Problem 6.3.8)

Let $\boldsymbol{A}$ be a symmetric matrix. If $\boldsymbol{u}$ and $\boldsymbol{v}$ are two eigenvectors of $\boldsymbol{A}$ associated with eigenvalues λ and μ, respectively, where $\lambda \neq \mu$, show that $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

Proof.

1. By assumption, we have $\boldsymbol{A}\boldsymbol{u} = \lambda \boldsymbol{u}$ (1) and $\boldsymbol{A}\boldsymbol{v} = \mu \boldsymbol{v}$ (2).
2. From Equation (1), we have $\boldsymbol{u}^T\boldsymbol{A} = \boldsymbol{u}^T\boldsymbol{A}^T = \lambda \boldsymbol{u}^T$. Hence $\boldsymbol{u}^T\boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{u}^T\boldsymbol{v}$.
3. From Equation (2), we have $\boldsymbol{u}^T\boldsymbol{A}\boldsymbol{v} = \boldsymbol{u}^T\mu\boldsymbol{v} = \mu\boldsymbol{u}^T\boldsymbol{v}$.
4. Therefore $\lambda\boldsymbol{u}^T\boldsymbol{v} = \mu\boldsymbol{u}^T\boldsymbol{v}$. Since $\lambda \neq \mu$, we have $\boldsymbol{u}^T\boldsymbol{v} = 0$, that is, $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

$\square$

Exercise

Let $\boldsymbol{A}$ be a symmetric matrix, λ an eigenvalue of $\boldsymbol{A}$, then λ is a real number.

Proof.

1. Assume $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$. Then $\boldsymbol{A}\bar{\boldsymbol{x}} = \bar{\boldsymbol{A}}\bar{\boldsymbol{x}} = \bar{\lambda}\bar{\boldsymbol{x}}$, that is, $\bar{\lambda}$ is an eigenvalue of $\boldsymbol{A}$.
2. Take transpose, we will get $\bar{\boldsymbol{x}}^T\boldsymbol{A} = \bar{\lambda}\bar{\boldsymbol{x}}$. Hence, $\bar{\boldsymbol{x}}^T\boldsymbol{A}\boldsymbol{x} = \bar{\lambda}\bar{\boldsymbol{x}}^T\boldsymbol{x}$.
3. Since $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, we have $\bar{\boldsymbol{x}}^T\boldsymbol{A}\boldsymbol{x} = \lambda\bar{\boldsymbol{x}}^T\boldsymbol{x}$.
4. Therefore $\bar{\lambda}\bar{\boldsymbol{x}}^T\boldsymbol{x} = \lambda\bar{\boldsymbol{x}}^T\boldsymbol{x}$. Since $\bar{\boldsymbol{x}}^T \neq 0$, $\lambda = \bar{\lambda}$, that is, λ is a real number.

$\square$

Exercise (5.33)

Determine which of the following statements are true. Justify your answer.

(a) *If $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are vectors in $\mathbb{R}^n$ such that $\boldsymbol{u}, \boldsymbol{v}$ are orthogonal and $\boldsymbol{v}, \boldsymbol{w}$ are orthogonal, then $\boldsymbol{u}, \boldsymbol{w}$ are orthogonal.*

(b) *If $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are vectors in $\mathbb{R}^n$ such that $\boldsymbol{u}, \boldsymbol{v}$ are orthogonal and $\boldsymbol{u}, \boldsymbol{w}$ are orthogonal, then $\boldsymbol{u}$ is orthogonal to $\operatorname{span}\{\boldsymbol{v}, \boldsymbol{w}\}$.*

(c) *If $\boldsymbol{A} = (\boldsymbol{c}_1 \;\; \boldsymbol{c}_2 \;\; \cdots \;\; \boldsymbol{c}_k)$ is an $n \times k$ matrix such that $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_k$ are orthonormal, then $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}_k$.*

(c') *If $\boldsymbol{A} = (\boldsymbol{c}_1 \;\; \boldsymbol{c}_2 \;\; \cdots \;\; \boldsymbol{c}_k)$ is an $n \times k$ matrix such that $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_k$ are orthogonal, then $\boldsymbol{A}^T\boldsymbol{A}$ is a diagonal matrix each of whose diagonal entries is not zero.*

(d) *If $\boldsymbol{A} = (\boldsymbol{c}_1 \;\; \boldsymbol{c}_2 \;\; \cdots \;\; \boldsymbol{c}_k)$ is an $n \times k$ matrix such that $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_k$ are orthonormal, then $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I}_n$.*

(e) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are orthogonal matrices, then $\boldsymbol{A} + \boldsymbol{B}$ is an orthogonal matrix.*

(f) *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are orthogonal matrices, then $\boldsymbol{AB}$ is an orthogonal matrix.*

(g) *If $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are the projections of $\boldsymbol{u}$ and $\boldsymbol{v}$ onto a vector space V, then $\boldsymbol{p}_1 + \boldsymbol{p}_2$ is the projection of $\boldsymbol{u} + \boldsymbol{v}$ onto V.*

(h) *If the columns of a square matrix $\boldsymbol{A}$ form an orthonormal set, then the rows of $\boldsymbol{A}$ also form an orthonormal set.*

(h') *If the columns of a square matrix $\boldsymbol{A}$ form an orthogonal set, then the rows of $\boldsymbol{A}$ also form an orthogonal set.*

Solution.

(a) False. For example: $\boldsymbol{u} = \boldsymbol{w} = (1, 0)$ and $\boldsymbol{v} = (0, 1)$.

(b) True. Let $a\boldsymbol{v} + b\boldsymbol{w}$ be any vector in $\mathrm{span}\{\boldsymbol{v}, \boldsymbol{w}\}$. Then $\boldsymbol{u} \cdot (a\boldsymbol{v} + b\boldsymbol{w}) = a(\boldsymbol{u} \cdot \boldsymbol{v}) + b(\boldsymbol{u} \cdot \boldsymbol{w}) = 0$.

(c) True. By definition.

(c') False. Take $\boldsymbol{c}_1$ to be $\mathbf{0}$.

(d) False. For example, let $\boldsymbol{A} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

(e) False. For example, let $\boldsymbol{A} = \boldsymbol{I}_2 = -\boldsymbol{B}$.

(f) True. $\boldsymbol{AB}(\boldsymbol{AB})^T = \boldsymbol{ABB}^T\boldsymbol{A}^T = \boldsymbol{AA}^T = \boldsymbol{I}$.

(g) True. By definition.

(h) True. By definition.

(h') False. For example, let $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

□

Exercise (Question 5(3-6) in Final 2005-2006(I))

Let $\boldsymbol{A}$ be an $n \times n$ matrix.

(3) *If $\boldsymbol{A}$ is diagonalizable and $\boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{x} = 0$ for every eigenvector $\boldsymbol{x}$ of $\boldsymbol{A}$, show that $\boldsymbol{A}$ is the zero matrix.*

(4) *Show that $\boldsymbol{B}\boldsymbol{B}^T + c\boldsymbol{I}$ is a symmetric matrix for any scalar c.*

(5) *Using the fact that any symmetric matrix is diagonalizable, prove that if $\|\boldsymbol{B}\boldsymbol{x}\| = \|\boldsymbol{x}\|$ for every $\boldsymbol{x} \in \mathbb{R}^n$, then $\boldsymbol{B}$ is an orthogonal matrix.*

(6) *We say that $\boldsymbol{C}$ preserves orthogonality if, for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$,*

$$\boldsymbol{x} \cdot \boldsymbol{y} = 0 \Rightarrow \boldsymbol{C}\boldsymbol{x} \cdot \boldsymbol{C}\boldsymbol{y} = 0.$$

Prove that if $\boldsymbol{C}$ preserves orthogonality, then $\boldsymbol{C}$ is a scalar multiple of an orthogonal matrix.

Exercise (Question 6 in Final 2001-2002(II))

Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a basis of $\mathbb{R}^n$ and let $\boldsymbol{A}$ be an $n \times n$ matrix. Prove that $\{\boldsymbol{A}\boldsymbol{v}_1, \boldsymbol{A}\boldsymbol{v}_2, \ldots, \boldsymbol{A}\boldsymbol{v}_n\}$ is a basis of $\mathbb{R}^n$ if and only if the nullspace of $\boldsymbol{A}$ is $\{\boldsymbol{0}\}$.

Change log

- Page 240: Revise a typo: "$(\boldsymbol{v})_{S_2}$" to "$(\boldsymbol{v})_{S_1}$";
- Page 244: Revise a mistake "general solution $\boldsymbol{x} = s(1,1,0)^T + t(-1,1,2)^T$" to "general solution $\boldsymbol{x} = s(1,1,0)^T + t(-1,0,1)^T$".

Last modified: 19:37, April 8, 2011.

Schedule of Tutorial 11

- Any question about last tutorial
- Review concepts: Linear transformation
  - Linear transformation, standard matrix;
  - Range, rank;
  - Kernal, nullity.
- Tutorial: 7.3, 7.5, 7.10, 7.11, 7.13, 7.14
- Additional material:
  - Question 3 in Final 2002–2003(II)
  - Question 5(d) in Final 2007–2008(II)
  - Question 3 in Final 2004–2005(II)

Linear Transformation

- A linear transformation is a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ of the form

$$T\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ for all } \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n,$$

  where a_{ij} is a real number for $1 \le i \le m$, $1 \le j \le n$. The matrix $(a_{ij})_{m\times n}$ is called the standard matrix for T.

- How to find the standard matrix for T:

$$\text{computing } \boldsymbol{A} = \begin{pmatrix} T(\boldsymbol{e}_1) & T(\boldsymbol{e}_2) & \cdots & T(\boldsymbol{e}_n) \end{pmatrix},$$

$$\text{or solving } T(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n) = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m)\boldsymbol{A}.$$

- Exercise 7.3: A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if and only if

$$T(a\boldsymbol{u} + b\boldsymbol{v}) = aT(\boldsymbol{u}) + bT(\boldsymbol{v}) \text{ for all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n, a, b \in \mathbb{R}.$$

  This result can be used in the Final Exam.

- If $n = m$, T is also called a linear operator on $\mathbb{R}^n$.

Linear Transformation (Cont.)

- Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, then
  (1) $T(\mathbf{0}) = \mathbf{0}$;
  (2) If $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k \in \mathbb{R}^n$ and $c_1, c_2, \ldots, c_k \in \mathbb{R}$, then
  $$T(c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k) = c_1 T(\boldsymbol{u}_1) + c_2 T(\boldsymbol{u}_2) + \cdots + c_k T(\boldsymbol{u}_k).$$
- Let $S : \mathbb{R}^n \to \mathbb{R}^m$ and $T : \mathbb{R}^m \to \mathbb{R}^k$ be linear transformations. The composition of T with S, denoted by $T \circ S$, is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^k$ such that
  $$(T \circ S)(\boldsymbol{u}) = T(S(\boldsymbol{u})) \quad \text{for all } \boldsymbol{u} \in \mathbb{R}^n.$$
- If $S : \mathbb{R}^n \to \mathbb{R}^m$ and $T : \mathbb{R}^m \to \mathbb{R}^k$ are linear transformations, then $T \circ S$ is again a linear transformation.
- If $\boldsymbol{A}$ and $\boldsymbol{B}$ are the standard matrices for the linear transformations S and T respectively, then the standard matrix for $T \circ S$ is $\boldsymbol{BA}$.

Range and Rank vs. Kernal and Nullity

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and $\boldsymbol{A}$ the standard matrix for T.

- - The range of T is the set of images of T, that is,
    $$\mathrm{R}(T) = \{ T(\boldsymbol{u}) \mid \boldsymbol{u} \in \mathbb{R}^n \} \subset \mathbb{R}^m.$$
  - $\mathrm{R}(T)$ is the column space of $\boldsymbol{A}$.
  - The dimension of $\mathrm{R}(T)$ is called the rank of T and denoted by $\mathrm{rank}(T)$.
  - $\mathrm{rank}(T) = \mathrm{rank}(\boldsymbol{A})$.
- - The kernal of T is the set of vectors in $\mathbb{R}^n$ whose image is the zero vector in $\mathbb{R}^m$, that is,
    $$\mathrm{Ker}(T) = \{\boldsymbol{u} \mid T(\boldsymbol{u}) = \boldsymbol{0}\} \subset \mathbb{R}^n.$$
  - $\mathrm{Ker}(T)$ is the nullspace of $\boldsymbol{A}$.
  - The dimension of $\mathrm{Ker}(T)$ is called the nullity of T and denoted by $\mathrm{nullity}(T)$.
  - $\mathrm{nullity}(T) = \mathrm{nullity}(\boldsymbol{A})$.
- $\mathrm{rank}(T) + \mathrm{nullity}(T) = \mathrm{rank}(\boldsymbol{A}) + \mathrm{nullity}(\boldsymbol{A}) = (\#$ columns of $\boldsymbol{A})$.
- For a general linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$, $\mathrm{Ker}(T) \in \mathbb{R}^n$ and $\mathrm{R}(T) \in \mathbb{R}^m$ are not necessarily in the same space.

Exercise (7.3)

Show that a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if and only if

$$T(a\boldsymbol{u} + b\boldsymbol{v}) = aT(\boldsymbol{u}) + bT(\boldsymbol{v}) \quad \textit{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n,\ a, b \in \mathbb{R}.$$

Proof.

"$\Rightarrow$" It is a particular case of Theorem 7.1.3.2.

"$\Leftarrow$"

1. Suppose $T(a\boldsymbol{u} + b\boldsymbol{v}) = aT(\boldsymbol{u}) + bT(\boldsymbol{v})$, for all $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, $a, b \in \mathbb{R}$.
2. Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ be the standard basis for $\mathbb{R}^n$ and $\boldsymbol{A}$ the $m \times n$ matrix
$$\begin{pmatrix} T(\boldsymbol{e}_1) & T(\boldsymbol{e}_2) & \cdots & T(\boldsymbol{e}_n) \end{pmatrix}.$$
We will see that $\boldsymbol{A}$ is the standard matrix for T:
3. For any $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)^T \in \mathbb{R}^n$, $\boldsymbol{u} = u_1\boldsymbol{e}_1 + u_2\boldsymbol{e}_2 + \cdots + u_n\boldsymbol{e}_n$, we have
$$T(\boldsymbol{u}) = u_1 T(\boldsymbol{e}_1) + u_2 T(\boldsymbol{e}_2) + \cdots + u_n T(\boldsymbol{e}_n) \qquad \text{By induction}$$
$$= \begin{pmatrix} T(\boldsymbol{e}_1) & T(\boldsymbol{e}_2) & \cdots & T(\boldsymbol{e}_n) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \boldsymbol{A}\boldsymbol{u}$$
4. Thus T is a linear transformation.

□

Exercise (7.5)

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator. If there exists a linear operator $S : \mathbb{R}^n \to \mathbb{R}^n$ such that $S \circ T$ is the identity transformation, i.e.

$$(S \circ T)(\boldsymbol{u}) = \boldsymbol{u} \quad \text{for all } \boldsymbol{u} \in \mathbb{R}^n,$$

then T is said to be the invertible and S is called the inverse of T.

(a) *For each of the following, determine whether T is invertible and find the inverse of T if possible.*

(i) $T : \mathbb{R}^2 \to \mathbb{R}^2$ *such that* $T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} y \\ x \end{pmatrix}$ *for all* $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$.

(ii) $T : \mathbb{R}^2 \to \mathbb{R}^2$ *such that* $T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x + y \\ 0 \end{pmatrix}$ *for all* $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$.

(b) *Suppose T is invertible and $\boldsymbol{A}$ is the standard matrix for T. Find the standard matrix for the inverse of T.*

Solution and Proof.

(a) (i) Since $T((x,y)^T) = (x,y)^T$ for all $(x,y)^T \in \mathbb{R}^2$, $(T \circ T)((x,y)^T) = T(T((x,y)^T)) = T((y,x)^T) = (x,y)^T$. That is, T is invertible, and its inverse is T itself.

(ii) Assume there exists an inverse $S : \mathbb{R}^2 \to \mathbb{R}^2$. Then $(1,0)^T = (S \circ T)((1,0)^T) = S((1,0)^T) = S \circ T((0,1)^T) = (0,1)^T$, a contradiction.

(b) The standard matrix of $S \circ T$ which is the product of the standard matrix of S and the standard matrix of T is identity matrix. That is,

$$\boldsymbol{BA} = \boldsymbol{I}_n$$

where $\boldsymbol{B}$ is the standard matrix of S. Hence the standard matrix of S is $\boldsymbol{A}^{-1}$.

□

Remark

- A linear operator T is invertible if and only if the standard matrix $\boldsymbol{A}$ of T is invertible. For part (a-ii), the standard matrix of T is $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, which is not invertible. Thus T is not invertible.
- A linear operator T is invertible if and only if it is bijective.

Exercise (7.10)

Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Show that $\mathrm{Ker}(T) = \{\mathbf{0}\}$ if and only if T is one-to-one, i.e. for any two vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, if $\boldsymbol{u} \neq \boldsymbol{v}$, then $T(\boldsymbol{u}) \neq T(\boldsymbol{v})$.

Proof.

"$\Rightarrow$"

1. Suppose that $\mathrm{Ker}(T) = \{\mathbf{0}\}$.
2. (Prove by contrapositive) Let $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ such that $T(\boldsymbol{u}) = T(\boldsymbol{v})$.
3. Then $T(\boldsymbol{u} - \boldsymbol{v}) = T(\boldsymbol{u}) - T(\boldsymbol{v}) = \mathbf{0}$, and hence $\boldsymbol{u} - \boldsymbol{v} \in \mathrm{Ker}(T)$.
4. Since $\mathrm{Ker}(T) = \{\mathbf{0}\}$, $\boldsymbol{u} - \boldsymbol{v} = \mathbf{0}$, that is, $\boldsymbol{u} = \boldsymbol{v}$. Thus T is one-to-one.

"$\Leftarrow$"

1. By Theorem 7.1.3.1, $T(\mathbf{0}) = \mathbf{0}$.
2. Since T is one-to-one, for all $\boldsymbol{v} \in \mathbb{R}^n$, if $\boldsymbol{v} \neq \mathbf{0}$, $T(\boldsymbol{v}) \neq T(\mathbf{0}) = \mathbf{0}$.
3. Thus $\mathrm{Ker}(T) = \{\mathbf{0}\}$.

$\square$

Exercise (7.11)

Let $S : \mathbb{R}^n \to \mathbb{R}^m$ and $T : \mathbb{R}^m \to \mathbb{R}^k$ be linear transformations.

(a) *Show that* $\mathrm{Ker}(S) \subset \mathrm{Ker}(T \circ S)$.

(b) *Show that* $\mathrm{R}(T \circ S) \subset \mathrm{R}(T)$.

Proof.

(a)
1. Let $\boldsymbol{u} \in \mathrm{Ker}(S)$, that is, $S(\boldsymbol{u}) = \mathbf{0}$.
2. Then $(T \circ S)(\boldsymbol{u}) = T(S(\boldsymbol{u})) = T(\mathbf{0}) = \mathbf{0}$ and hence $\boldsymbol{u} \in \mathrm{Ker}(T \circ S)$.
3. Thus $\mathrm{Ker}(S) \subset \mathrm{Ker}(T \circ S)$.

(b)
1. Let $\boldsymbol{v} \in \mathrm{R}(T \circ S)$, that is, there exists $\boldsymbol{u} \in \mathbb{R}^n$ such that $\boldsymbol{v} = (T \circ S)(\boldsymbol{u})$.
2. Put $\boldsymbol{w} = S(\boldsymbol{u}) \in \mathbb{R}^m$. Then $\boldsymbol{v} = T(S(\boldsymbol{u})) = T(\boldsymbol{w})$.
3. This means that $\boldsymbol{v} \in \mathrm{R}(T)$. Thus $\mathrm{R}(T \circ S) \subset \mathrm{R}(T)$.

□

Exercise (7.13(a))

Let $\boldsymbol{n}$ be a unit vector in $\mathbb{R}^n$. Define $F : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$F(\boldsymbol{x}) = \boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} \quad \textit{for all } \boldsymbol{x} \in \mathbb{R}^n.$$

(a) *Show that F is a linear transformation and find the standard matrix for F.*

Proof of part (a).

For any $\boldsymbol{x} \in \mathbb{R}^n$,

$$\begin{aligned} F(\boldsymbol{x}) &= \boldsymbol{x} - 2\boldsymbol{n}(\boldsymbol{n} \cdot \boldsymbol{x}) && \boldsymbol{n} \cdot \boldsymbol{x} \text{ is a real number.} \\ &= \boldsymbol{x} - 2\boldsymbol{n}(\boldsymbol{n}^T\boldsymbol{x}) && \text{Definition of inner product.} \\ &= \boldsymbol{x} - 2(\boldsymbol{n}\boldsymbol{n}^T)\boldsymbol{x} && \text{Associated law of matrix product.} \\ &= (\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T)\boldsymbol{x} \end{aligned}$$

So F is a linear transformation, whose standard matrix is $\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T$. □

Exercise (7.13(b))

Let $\boldsymbol{n}$ be a unit vector in $\mathbb{R}^n$. Define $F : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$F(\boldsymbol{x}) = \boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} \quad \textit{for all } \boldsymbol{x} \in \mathbb{R}^n.$$

(b) *Show that $F \circ F$ is the identity transformation.*

Proof of part (b).

- (First method) For any $\boldsymbol{x} \in \mathbb{R}^n$, we have

$$\begin{aligned}
&(F \circ F)(\boldsymbol{x}) = F(F(\boldsymbol{x})) = F(\boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n}) && \text{Definition of } F.\\
&= \boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} - 2\{\boldsymbol{n} \cdot [\boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n}]\}\boldsymbol{n} && \text{Definition of } F.\\
&= \boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} - 2\{(\boldsymbol{n} \cdot \boldsymbol{x}) - 2(\boldsymbol{n} \cdot \boldsymbol{x})(\boldsymbol{n} \cdot \boldsymbol{n})\}\boldsymbol{n} && \text{Distributive law of inner product.}\\
&= \boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} - 2\{-(\boldsymbol{n} \cdot \boldsymbol{x})\} \cdot \boldsymbol{n} = \boldsymbol{x} && \boldsymbol{n} \text{ is a unit vector.}
\end{aligned}$$

  Therefore, $F \circ F$ is the identity transformation.

- (Second method) Alternatively, we consider the standard matrix of $F \circ F$:

$$(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T)^2 = (\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T)(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T) = \boldsymbol{I} - 4\boldsymbol{n}\boldsymbol{n}^T + 4\boldsymbol{n}\boldsymbol{n}^T\boldsymbol{n}\boldsymbol{n}^T = \boldsymbol{I}.$$

  Therefore, $F \circ F$ is the identity transformation.

$\square$

Exercise (7.13(c))

Let $\boldsymbol{n}$ be a unit vector in $\mathbb{R}^n$. Define $F : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$F(\boldsymbol{x}) = \boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} \quad \textit{for all } \boldsymbol{x} \in \mathbb{R}^n.$$

(c) *Show that the standard matrix for F is an orthogonal matrix.*

Proof of part (c).

1. By part (a), the standard matrix of F is $\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T$.
2. By part (b), $(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T)^{-1} = \boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T$.
3. Note that $(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T)^T = \boldsymbol{I} - 2(\boldsymbol{n}\boldsymbol{n}^T)^T = \boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T$.
4. Thus
$$(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T)^T = (\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T)^{-1},$$
that is $\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T$ is an orthogonal matrix.

□

Remark

F is a reflection operator about the hyperplane which is orthogonal to $\boldsymbol{n}$.

Exercise (7.14)

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation such that $T \circ T = T$.

(a) *If T is not the zero transformation, show that there exists a nonzero vector $\boldsymbol{u} \in \mathbb{R}^n$ such that $T(\boldsymbol{u}) = \boldsymbol{u}$.*

(b) *If T is not the identity transformation, show that there exists a nonzero vector $\boldsymbol{v} \in \mathbb{R}^n$ such that $T(\boldsymbol{v}) = \boldsymbol{0}$.*

(c) *Find all linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $T \circ T = T$.*

Proof of parts (a) and (b).

(a) Suppose T is not the zero transformation. So there exists $\boldsymbol{x} \in \mathbb{R}^n$ such that $T(\boldsymbol{x}) \neq 0$. Define $\boldsymbol{u} = T(\boldsymbol{x})$. Then $\boldsymbol{u}$ is a nonzero vector and

$$T(\boldsymbol{u}) \xlongequal{\boldsymbol{u}=T(\boldsymbol{x})} T(T(\boldsymbol{x})) = (T \circ T)(\boldsymbol{x}) \xlongequal{T \circ T = T} T(\boldsymbol{x}) \xlongequal{\boldsymbol{u}=T(\boldsymbol{x})} \boldsymbol{u}.$$

(b) Suppose T is not the identity transformation. So there exists $\boldsymbol{y} \in \mathbb{R}^n$ such that $T(\boldsymbol{y}) \neq \boldsymbol{y}$. Define $\boldsymbol{v} = T(\boldsymbol{y}) - \boldsymbol{y}$. Then $\boldsymbol{v}$ is a nonzero vector and

$$T(\boldsymbol{v}) = T(T(\boldsymbol{y}) - \boldsymbol{y}) = (T \circ T)(\boldsymbol{y}) - T(\boldsymbol{y}) = T(\boldsymbol{y}) - T(\boldsymbol{y}) = \boldsymbol{0}.$$

$\square$

Solution of part(c).

- Let $\boldsymbol{A}$ be the standard matrix for T. Then it is equivalent to find all 2×2 matrices $\boldsymbol{A}$, such that $\boldsymbol{A}^2 = \boldsymbol{A}$.
- Let λ be an eigenvalue of $\boldsymbol{A}$, and $\boldsymbol{x}$ an eigenvector associated with λ, then $\lambda^2 \boldsymbol{x} = \boldsymbol{A}^2 \boldsymbol{x} = \boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$. Since $\boldsymbol{x}$ is nonzero vector, $\lambda^2 = \lambda$. Hence λ can only be 0 or 1.
- Case 1: $\lambda_1 = \lambda_2 = 0$. We cannot find a nonzero vector $\boldsymbol{u}$, such that $T(\boldsymbol{u}) = \boldsymbol{u}$; Otherwise T has an eigenvalue 1. By part (a), then T is the zero transformation.
- Case 2: $\lambda_1 = \lambda_2 = 1$. We cannot find a nonzero vector $\boldsymbol{v}$, such that $T(\boldsymbol{v}) = \boldsymbol{0}$; Otherwise T has an eigenvalue 0. By part (b), then T is the identity transformation.
- Case 3: $\lambda_1 = 0$, $\lambda_2 = 1$. Then $\boldsymbol{A}$ can be diagonalizable. Then $\boldsymbol{A} = \boldsymbol{P}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \boldsymbol{P}$ for some invertible matrix $\boldsymbol{P}$. Let $\boldsymbol{P} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $\boldsymbol{A} = \frac{1}{ad-bc} \begin{pmatrix} ad & bd \\ -ac & -bc \end{pmatrix}$ where $ad - bc \neq 0$. We can simplify the expression to $\begin{pmatrix} r & s \\ t & 1-r \end{pmatrix}$ where $st = r(1-r)$.
- Therefore
$$\boldsymbol{A} = \boldsymbol{0}_2, \boldsymbol{I}_2, \begin{pmatrix} r & s \\ t & 1-r \end{pmatrix}, \text{ where } st = r(1-r).$$

$\square$

Exercise (Question 3 in Final 2002-2003(II))

Let $V = \mathrm{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ *where* $\boldsymbol{u}_1 = (1, 2, 3)$ *and* $\boldsymbol{u}_2 = (1, 1, 1)$*.*

(a) *Find all vectors orthogonal to* V*.*

(b) *Note that* V *is a plane in* $\mathbb{R}^3$ *containing the origin. Write down an equation that represents this plane.*

Exercise (Question 5(d) in Final 2007-2008(II))

Determine whether the statements is true: If the nullspace of two matrices $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *are the same, then* $\boldsymbol{A}$ *is row equivalent to* $\boldsymbol{B}$*.*

Exercise (Question 3 in Final 2004-2005(II))

Let $\{e_1, e_2, \ldots, e_n\}$ be the basis of $\mathbb{R}^n$ and let T be a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$ such that $T(\boldsymbol{e}_i) = \boldsymbol{e}_{i+1}$ for $i = 1, 2, \ldots, n-1$ and $T(\boldsymbol{e}_n) = \mathbf{0}$. Find all the eigenvalues and eigenvectors of $\boldsymbol{A}$, where $\boldsymbol{A}$ is the standard matrix for T.

Solution.

$$T(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) = (\boldsymbol{e}_2, \ldots, \boldsymbol{e}_n, \mathbf{0}) = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) \begin{pmatrix} 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 1 & 0 \end{pmatrix}.$$

□

Exercise (Question 3(b) in Final 2005-2006(I))

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. If $T \circ T = T$, show that

$$\mathrm{Ker}(T) \cap \mathrm{R}(T) = \{\mathbf{0}\}.$$

Change log

- Page 258: Revise a typo: "$\begin{pmatrix} x \\ y \end{pmatrix}$" to "$\begin{pmatrix} y \\ x \end{pmatrix}$";
- Page 259: Revise a mistake for part (a-i).

Last modified: 20:15, April 15, 2011.

Thank you