MA1101R Tutorial

Xiang Sun[1,2]

Dept. Mathematics

April 13, 2010

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

[2] Corrections are always welcome.

Self-Introduction

Name Sun Xiang (English) and 孙祥 (Chinese)
Second year Ph.D. student in Dept. Mathematics

Email xiangsun@nus.edu.sg

Mobile 9053 5550

Office S17-06-14

IM xiangsun.sunny@hotmail.com (MSN), 402197754 (QQ)

Introduction

- 11 tutorials: 4 before mid-term test, and 7 after it;
- Take attendance:
  - 2 points for full attendance, and pro-rated for partial attendance;
  - Everyone need to print his/her signature, rather than just a tick;
  - If you find some mistakes on the attendance sheet, please let me know.
- Presentation:
  - 1 point for at least 1 presentation for tutorial problems;
  - I will call name one-by-one or call for volunteers from next tutorial.
- Tutorial style:
  - 5–10 mins for review;
  - 35–45 mins for tutorial questions;
  - 0–10 mins for additional material.
- Additional material: discuss questions in the past-year papers, some anecdotes and histories.
- Download: click here to my SkyDrive.

Schedule of Today

- Review concepts
- Tutorial: 1.9, 1.15, 1.18(c), 1.19, 1.22, 1.23
- Additional material

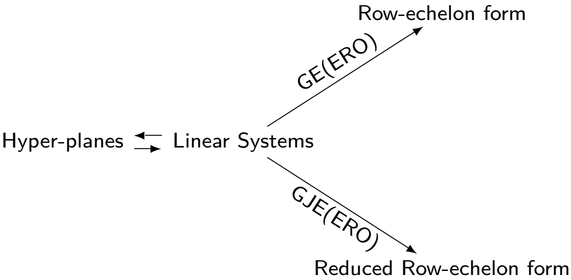
Linear Algebra

Linear algebra is the branch of mathematics devoted to the theory of linear structure (vector space, module), representation of the structure (vector space associated linear transformation) and some relative issue. Matrix theory is one of the most important tools of linear algebra.

- Chapter 1 is an introduction for linear algebra;
- In chapter 2,4,6, we will discuss matrix theory;
- In chapter 3,5, we will discuss vector space and a special vector space—Euclidean space;
- In chapter 7, we will discuss linear transformation.

History of Linear System

- 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, 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 published the rule (Thm 2.5.32) named after him for the solution of an $n \times n$ system;
- Euler 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 introduced a systematic procedure, now called Gaussian elimination, for the solution of systems of linear equations, though he did not use the matrix notation.


Row-echelon form
GE(ERO)
Hyper-planes
Linear Systems
GJE(ERO)
Reduced Row-echelon form

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 $\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$

Row-Echelon Form and Reduced Row-Echelon Form

- If the augmented matrix of a linear system is in REF or RREF, we can get the solutions easily.
- An augmented matrix is said to be in row-echelon form(REF) if it has properties 1, 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 non-zero rows, the first nonzero number in the lower row occurs farther to the right than the first nonzero number in the higher row.
-
  - 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.
  - In a REF, (# of non-zero rows) = (# of leading entries) = (# of pivot columns) = (# of pivot points).
- An augmented matrix is said to be in reduced row-echelon form(RREF) if it is has properties 1, 2, 3, 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.
-
  - 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.
  - A linear system has exactly one solution if except the last column, every column of a REF of the augmented matrix is a pivot column, i.e. (# of variables) = (# of nonzero rows).
  - A linear system has infinitely many solutions if apart from the last column, a REF of the augmented matrix has at least one more non-pivot column, i.e. (# of variables) $>$ (# of nonzero rows). In this case, its general solution has (# of variables $-$ # of nonzero rows) arbitrary parameter(s).

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

- There are some 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.

Exercise (1.9)

Each equation in the following linear system represents a plane in the xyz-space

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{cases}$$

where a_1, a_2, b_1, b_2, c_1, c_2, d_1, d_2 are constants. Discuss the relative positions of the two planes when:

(a) *has no solution;*

(b) *has exactly one solution;*

(c) *has infinitely many solutions and a general solution has one arbitrary parameter;*

(d) *has infinitely many solutions and a general solution has two arbitrary parameters.*

Solution.

- From geometric view, we can see that there are 2 cases of relative positions of two planes, and 2 subcases of the parallel case:

$$\begin{cases} \text{parallel:} & \begin{cases} \text{same—infinitely many solutions and a general solution has two arbitrary parameters} \\ \text{distinct—no solution} \end{cases} \\ \text{nonparallel:} & \text{intersection is a line—infinitely many solutions and a general solution has one arbitrary parameter} \end{cases}$$

- There are 3 variables and 2 equations, so (# of variables) $= 3 > 2 \geq$ (# of nonzero rows), i.e., the linear system can not have exactly one solution.

To summarize:

(a) Two planes are parallel but distinct.

(b) Such a situation does not exist.

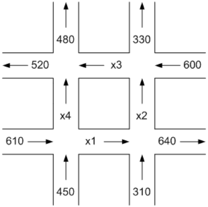
(c) Two planes are not parallel.

(d) Two planes are identical.

$\square$

Exercise (1.15)

In the downtown section of a certain city, two sets of one-way streets intersect as shown in the following:



The average hourly volume of traffic entering and leaving this section during rush hour is given in the diagram.

(a) *Do we have enough information to find the traffic volumes* x_1, x_2, x_3, x_4? *Explain your answer.*

(b) *Given that* $x_4 = 500$, *find* x_1, x_2, x_3.

(The average hourly volume of traffic entering an intersection must be equal to the volume that leaving.)

Solution.

(a) At each traffic junction, use the incoming and outgoing arrow to form a linear equation. Therefore we have:

$$\begin{cases} x_1 + 310 = 640 + x_2 \\ 610 + 450 = x_1 + x_4 \\ 600 + x_2 = x_3 + 330 \\ x_3 + x_4 = 520 + 480 \end{cases} \Leftrightarrow \begin{cases} x_1 - x_2 = 330 \\ x_1 + x_4 = 1060 \\ x_2 - x_3 = -270 \\ x_3 + x_4 = 1000 \end{cases} \quad (*)$$

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 330 \\ 1 & 0 & 0 & 1 & 1060 \\ 0 & 1 & -1 & 0 & -270 \\ 0 & 0 & 1 & 1 & 1000 \end{array}\right) \xrightarrow{\text{Gaussian Elimination}} \left(\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 330 \\ 0 & 1 & 0 & 1 & 730 \\ 0 & 0 & -1 & -1 & -1000 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Therefore, the system has infinitely many solutions. We cannot determine the values of x_1, x_2, x_3, x_4 uniquely.

(b) By solving the equation $(*)$, we have $x_1 = 560$, $x_2 = 230$, $x_3 = 500$. $\square$

Exercise (1.18(c))

For

$$\begin{cases} ax + ay & = 1 \\ ax - |a|y & = 1 \end{cases}$$

determine the values of a such that the system has

(i) *no solution;*

(ii) *exactly one solution;*

(iii) *infinitely many solutions.*

Solution.

There is some difficulty need to deal with: $|a|$. So we need to consider by cases:

$a = 0$ It is obvious that there is no solution;

$a > 0$ The system becomes $\begin{cases} ax + ay & = 1 \\ ax - ay & = 1 \end{cases}$, so there is unique solution $\begin{cases} x = 1/a \\ y = 0 \end{cases}$;

$a < 0$ The system becomes $\begin{cases} ax + ay & = 1 \\ ax + ay & = 1 \end{cases} \Leftrightarrow ax + ay = 1$, so the system has infinitely many solutions and a general solution has one arbitrary parameter;

To summarize: The system has no solution if $a = 0$, only one solution if $a > 0$, infinitely many solutions if $a < 0$. $\square$

Exercise (1.19)

Determine the values of a and b so that the system

$$\begin{cases} ax + bz = 2 \\ ax + ay + 4z = 4 \\ ay + 2z = b \end{cases}$$

(a) *has no solution;*

(b) *has exactly one solution;*

(c) *has infinitely many solutions and a general solution has one arbitrary parameter;*

(d) *has infinitely many solutions and a general solution has two arbitrary parameters.*

Solution.

Perform elementary row operations on the augmented matrix:

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

Consider two cases: $a \neq 0$, and $a = 0$.

$a \neq 0$: the first two columns of the last augmented matrix are pivot columns. So, if $b-2 \neq 0$, then the third column will also be pivotal, and the system has a unique solution. On the other hand, if $b-2=0$, then the system has infinitely many solutions with one parameter.

$a = 0$: the last augmented matrix becomes $\left(\begin{array}{ccc|c} 0 & 0 & b & 2 \\ 0 & 0 & 4-b & 2 \\ 0 & 0 & b-2 & b-2 \end{array}\right)$.

In view of row 1 and 2, if $b \neq 4-b$, then these two rows will not be consistent, and hence the system will have no solution. If $b = 4-b$, then $b=2$ and we have a consistent system. In this case, since there are only one pivot columns, the system has infinitely many solutions with two parameters.

To summarize:

(a) no solution: $a = 0$ and $b \neq 2$.

(b) only one solution: $a \neq 0$ and $b \neq 2$.

(c) infinitely many solutions with one arbitrary parameter: $a \neq 0$ and $b = 2$.

(d) infinitely many solutions with two arbitrary parameters: $a = 0$ and $b = 2$.

□

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 must have that exactly two of a, b, c are zeros. Since it is given that the matrix is in reduced row-echelon form, a can not be 0. Therefore $a = 1, b = c = 0$.
- Since the system is consistent (Solution set is a plane.) and $b = c = 0$, we have $e = f = 0$.
- Since the plane does not contain the origin, $d \neq 0$.

To summarize: the sufficient 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 more 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: $\begin{cases} x+y &= 0 \\ x-y &= 0 \\ 2x+4y &= 0 \end{cases}$.

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

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

Necessary and Sufficient condition

- The rank of a matrix(pp.119 of textbook) is the dimension of its row space (or column space). Notation: $\mathrm{rank}(A)$. $\mathrm{rank}(A)$ is equal to the number of nonzero pivot columns in a REF of A.
- Homogeneous:
  - $A_{m\times n}\cdot x_{n\times 1}=0_{m\times 1}, \mathrm{rank}(A)=r\leq \min\{m,n\}$;
  - The system can be always solved;
  - If $r=n$, then there is only zero solution;
  - If $r<n$, then there are infinite solutions with $n-r$ arbitrary parameter(s).
- Inhomogeneous:
  - $A_{m\times n}\cdot x_{n\times 1}=b_{m\times 1}, \mathrm{rank}(A)=r\leq \min\{m,n\}$;
  - The system can be solved iff $\mathrm{rank}(A)=\mathrm{rank}(A\ \ b)$.
  - 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).

For a linear system, can all the coefficients be zeros?

Page 7 Add a relation diagram;

Page 15 For part (b): “By part (a)”–>“By solving the equation (*)”;

Page 18 The second row: “Perform Gaussian elimination”–>“Perform elementary row operations”;

Page 22 Add a question which is related with Exercise 1.21.

Last modified: 16:44, January 26, 2010.

Schedule of Today

- Any question about last tutorial
- Review concepts
- Tutorial: 2.9, 2.14, 2.16, 2.20, 2.22, 2.27
- Additional material

Remarks on Reference books

- Hoffman K, Kunze R, *Linear Algebra (2nd Edition)*, Prentice Hall, Inc., New Jersey, 1971
- Roman S, *Advanced Linear Algebra (2nd Edition)*, Springer, New York, 2005
- Jacobson N, *Lectures in Abstract Algebra II: Linear Algebra*, Springer, New York, 1953
- 李炯生，查建国，线性代数，中国科学技术大学出版社，合肥，1989
- 许以超，线性代数与矩阵论（第二版），高等教育出版社，北京，2008
- 张贤科，徐甫华，高等代数学（第二版），清华大学出版社，北京，2004
- 李尚志，线性代数，高等教育出版社，北京，2006
- 龚升，线性代数五讲，中国科学技术大学出版社，合肥，2005

History of Matrix Theory

- Matrices were introduced implicitly as abbreviations of linear transformations by Gauss;
- Arthur Cayley formally introduced $m \times n$ matrices in two papers in 1850 and 1858 (the term "matrix"was coined by Sylvester 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 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, 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 developed substantial elements of the theory of matrices in the language of bilinear forms.

Definitions and Notations

- 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 # of rows and n is # of 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

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

  or simply $A = (a_{ij})_{m \times n}$ where a_{ij} is the (i,j)-entry of A.

Special Matrices I

Row matrix only 1 column;

Column matrix only 1 row;

Square matrix # of rows = # of 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;

Upper-tridiagonal matrix nonzero elements only in the main diagonal, the first diagonal above this;

- * Multiplication of two upper-triangular matrices is also an upper-triangular matrix;
- * Inverse of an upper-triangular invertible matrix is upper-triangular;

Lower-tridiagonal matrix nonzero elements only in the main diagonal, the first diagonal below this;

Identity matrix diagonal matrix where diagonal entries are 1, notation: I_n;

Scalar matrix diagonal matrix where diagonal entries are c–constant number, cI_n;

Zero matrix all entries are 0, notation: $0_{m\times n}$;

Addition

- Two matrices are said to be equal if they have the same size and their corresponding entries are equal.
- Let $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{m\times n}$. Define the matrix addition

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

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

Scalar Multiplication

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

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

Multiplication

Let $A = (a_{ij})_{m\times p}$ and $B = (b_{ij})_{p\times n}$. Define the matrix multiplication $AB = (c_{ij})_{m\times n}$, 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 A, B and C be $m\times p$, $p\times q$ and $q\times n$ matrices respectively, then $(AB)C = A(BC)$; Moreover, we can define $A^n = \begin{cases} I & \text{if } n = 0 \\ \overbrace{AA\cdots A}^{n \text{ times}} & \text{if } n \geq 1 \end{cases}$.
- Commutative law: not always hold.
- Identity: Let $A = (a_{ij})_{m\times n}$, then $AI_n = I_mA = A$;
- Inverse: not invertible for all matrices;
- 1st-Distributive law: Let A, B_1 and B_2 be $m\times p$, $p\times n$ and $p\times n$ matrices respectively, then $A(B_1 + B_2) = AB_1 + AB_2$;
- 2nd-Distributive law: Let A, C_1 and C_2 be $p\times n$, $m\times p$ and $m\times p$ matrices respectively, then $(C_1 + C_2)A = C_1A + C_2A$;
- Let $A = (a_{ij})_{m\times p}$, $B = (b_{ij})_{p\times n}$ and μ be a real constant, then $(\mu A)B = A(\mu B) = \mu(AB)$;

Transpose

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

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

Inverse

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

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

Special Matrices II

Symmetric matrix $A = A^T$;

Skew-symmetric matrix $A = -A^T$;

Hermite matrix $A = \bar{A}^T$;

* Let A be a square matrix, then $A + A^T$ is a symmetric matrix, and $A - A^T$ is a skew-symmetric matrix;
* Each square matrix A can be uniquely decomposed as an addition of symmetric matrix S and a skew-symmetric matrix K.

Nilpotent matrix $A^k = 0$ for some positive integer k;

* Let A be a matrix with $A^k = 0$, then $(I - A)^{-1} = I + A + \cdots + A^{k-1}$.

Idempotent matrix $A^2 = A$;

* Let A be an idempotent matrix, then $(I + A)^{-1} = \frac{1}{2}(2I - A)$.

Elementary Matrices

1. Multiply a row by a constant:

$$E = \begin{array}{cc} & \begin{array}{ccccccc} & & & i\text{th} & & & \end{array} \\ i\text{th} & \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & & & & \\ & & & c & & & \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{array}, \quad E^{-1} = \begin{array}{cc} & \begin{array}{ccccccc} & & & i\text{th} & & & \end{array} \\ i\text{th} & \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & & & & \\ & & & \frac{1}{c} & & & \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{array}$$

Elementary Matrices

2. Interchange two rows:

$$E = \begin{array}{cc}
 & \begin{array}{ccccccccccc} & & & i\text{th} & & & & j\text{th} & & & \end{array} \\
\begin{array}{c} \\ \\ \\ i\text{th} \\ \\ \\ \\ j\text{th} \\ \\ \\ \\ \end{array} &
\begin{pmatrix}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & 0 & \\
 & & 1 & & & & & & & & \\
 & & & {\color{red}0} & & & & {\color{red}1} & & & \\
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
 & & & {\color{red}1} & & & & {\color{red}0} & & & \\
 & & & & & & & & 1 & & \\
 & 0 & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{pmatrix}
\end{array} = E^{-1}$$

Elementary Matrices

3. Add a multiple of a row by a constant:

$$E = \begin{array}{cc} & \begin{array}{ccccccccc} & & & i\text{th} & & j\text{th} & & & \end{array} \\ \begin{array}{c} \\ \\ \\ i\text{th} \\ \\ j\text{th} \\ \\ \\ \\ \end{array} & \begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&& 0 & \\ && 1 &&&&&& \\ &&& 1 && c &&& \\ &&&& \ddots &&&& \\ &&&&& 1 &&& \\ &&&&&& 1 && \\ &&&&&&& \ddots & \\ &&&&&&&& 1 \end{pmatrix} \end{array}, \text{ and } E = \begin{array}{cc} & \begin{array}{ccccccccc} & & & j\text{th} & & i\text{th} & & & \end{array} \\ \begin{array}{c} \\ \\ \\ j\text{th} \\ \\ i\text{th} \\ \\ \\ \\ \end{array} & \begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&& 0 & \\ && 1 &&&&&& \\ &&& 1 &&&&& \\ &&&& \ddots &&&& \\ &&& c && 1 &&& \\ &&&&&& 1 && \\ &&&&&&& \ddots & \\ &&&&&&&& 1 \end{pmatrix} \end{array}$$

Exercise (2.9)

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

Proof.

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

$$A(u + v) = Au + Av = b + 0 = b.$$

Hence $Ax = b$ has either no solutions or infinitely many solutions. $\square$

Remark

The structure of the solution set of inhomogeneous system $Ax = b$:

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

where u is any specific solution to the linear system $Ax = b$.

2nd Method.

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

$\square$

Exercise (2.14)

Let $\mathcal{M}$ *be the set of all* 2×2 *matrices. Let* $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ *and* $\mathcal{B} = \{X \in \mathcal{M} \mid AX = XA\}$.

(a) *Determine which of the following elements of* $\mathcal{M}$ *lies in* $\mathcal{B}$*:*

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

(b) *Prove that if* $P, Q \in \mathcal{B}$*, then* $P + Q \in \mathcal{B}$ *and* $PQ \in \mathcal{B}$.

(c) *Find conditions on* p, q, r, s *which determine precisely when* $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ *lies in* $\mathcal{B}$.

Solution and Proof.

(a) Substituting each matrix for X in $AX = XA$, and then we can see whether the equation is satisfied: All except $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ lies in B.

(b) Since $P, Q \in \mathcal{B}$, we have $AP = PA$ and $AQ = QA$. Then

$$A(P+Q) = AP + AQ = PA + QA = (P+Q)A,$$

$$A(PQ) = (AP)Q = (PA)Q = P(AQ) = P(QA) = (PQ)A,$$

where we apply distribution law and associated law in 1st equation and 2nd equation, respectively.

(c) When $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ lies in $\mathcal{B}$, then p, q, r, s must satisfy

$$\begin{pmatrix} p+r & q+s \\ r & s \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p & q \\ r & s \end{pmatrix} = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} p & p+q \\ r & r+s \end{pmatrix}.$$

Hence we have 4 equations about p, q, r, s. Solving the equations we will get that $p = s$, $r = 0$, and q is arbitrary. That is to say, when $p = s$ and $r = 0$, the matrix $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ lies in $\mathcal{B}$.

□

Exercise (2.16)

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

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

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

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

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

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

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

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

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

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

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

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

(d) True. $(A-B)^T = A^T - B^T = A - B$

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

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

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

(g) True. Let $A = (a_{ij})_{n\times n}$, then for each $i \in \{1, 2, \dots, n\}$, (i, i)-entry of 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 $AA^T = 0$ implies that $a_{ik} = 0$ for all i and k, i.e. $A = 0$.

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

Remark

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

Exercise (2.20)

Let A be a square matrix.

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

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

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

Recall

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

Proof and Solution.

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

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

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

$\square$

Exercise (2.22)

Let A and B be invertible matrices of the same size.

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

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

Solution and Proof.

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

(b)

$$\begin{aligned} A^{-1} + B^{-1} &= B^{-1}(BA^{-1} + I) \\ &= B^{-1}(B + A)A^{-1} \end{aligned}$$

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

$\square$

Exercise (2.27)

Let A, B be 3×3 matrices such that $A = E_1 E_2 E_3 E_4 B$ where

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(a) *Describe how A is obtained from B by elementary row operations.*

(b) *If A is invertible, is B invertible? Justify your answer.*

Solution.

(a) $B \xrightarrow{R_1 - R_3} \xrightarrow{R_1 \leftrightarrow R_3} \xrightarrow{R_3 + 2R_2} \xrightarrow{2R_3} A$

(b) Since $B = E_4^{-1} E_3^{-1} E_2^{-1} E_1^{-1} A$ and $E_1^{-1}, E_2^{-1}, E_3^{-1}, E_4^{-1}$ are invertible, if A is invertible, then B is invertible.

$\square$

Exercise (2.11)

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

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

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

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

Exercise (2.15)

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

Exercise

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

Solution.

- $$A\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} = \begin{pmatrix}2\\ \vdots\\ 2\end{pmatrix}, \quad A^2\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} = 2A\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} = \begin{pmatrix}4\\ \vdots\\ 4\end{pmatrix}, \quad 2J\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} = \begin{pmatrix}2n\\ \vdots\\ 2n\end{pmatrix}$$

  Hence $4 + 4 = 2n$, i.e. $n = 4$, A is a matrix of order 4.
- If B satisfies $A^2 + 2A = 2J$, so does B^T.
- The task left is simple.

□

Exercise (Question 1)

Given an invertible matrix, how to compute its inverse.

Exercise (Question 2)

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

For Exercise 1.21, the solution is not correct: there is another RREF

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

- **Page 6** Change "Upper-triangle" and "Lower-triangle" to "Upper-tridiagonal" and "Lower-tridiagonal", respectively;
- **Page 17** Add another method of Exercise 2.9;
- **Page 22** Add another example for part (f) and a remark;
- **Page 23** Change "$(I - A)(I + A \cdots + A^n) = I - A^n$" to "$(I - A)(I + A \cdots + A^{n-1}) = I - A^n$";
- **Page 23** Delete a remark;
- **Page 24** Rewrite the proof of part (b).

Last modified: 10:49, March 21, 2010.

Schedule of Today

- Any question about last tutorial
- Review concepts
- Tutorial: 2.36, 2.37, 2.40(b), 2.46, 2.48, 2.49
- Additional material

1st-definition of Determinant (Laplace formula)

- Let $A = (a_{ij})$ be an $n \times n$ matrix. Let M_{ij} be an matrix obtained from A by deleting the i-th row and the j-th column. Then the determinant of A is defined as

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

  where $A_{ij} = (-1)^{i+j}\det(M_{ij})$, which is called the (i, j)-cofactor of A.
- Let $A = (a_{ij})$ be an $n \times n$ matrix. $\det(A)$ is usually written as

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

2nd-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

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

3rd-definition of Determinant (Axioms)

- Let D be a function from $n \times n$ 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.
- Let D be a function from $n \times n$ matrices to $\mathbb{R}$. We say that D is alternating if the following two conditions are satisfied:
  - $D(A) = 0$ whenever two rows of A are equal;
  - If A' is a matrix obtained from A by interchanging two rows of A, then $D(A') = -D(A)$.
- Let D be a function from $n \times n$ matrices to $\mathbb{R}$. We say that D is a determinant function if D is n-linear, alternating, and $D(I_n) = 1$.
- Existence and uniqueness: Corollary, page 147 and Theorem 2, page 152 of Hoffman's "Linear Algebra". Notation: $\det$.

Equivalence of the 3 definitions

- Def 3 $\Rightarrow$ Def 1 Theorem 1, page 146 of Hoffman's "Linear Algebra";
- Def 1 $\Rightarrow$ Def 3 Trivial;
- Def 2 $\Rightarrow$ Def 1 Section 2.3, 许以超的 "线性代数与矩阵论"; Moreover, we will get Laplace Expansions (Ref Theorem 2.3.3 of 许以超的 "线性代数与矩阵论" or section 5.7 of Hoffman's "Linear Algebra");
- Def 1 $\Rightarrow$ Def 2 Mathematical Induction.

Properties of Determinants

- If A is a square matrix, then $\det(A) = \det(A^T)$;
- The determinant of a square matrix with two identical rows (columns) is zero;
- Let A be a square matrix.
  - If B is obtained from A by multiplying one row of A by a constant k, then $\det(B) = k\det(A)$;
  - If B is obtained from A by interchanging two rows of A, then $\det(B) = -\det(A)$;
  - If B is obtained from A by adding a multiple of one row of A to another row, then $\det(B) = \det(A)$.
- A square matrix A is invertible if and only if $\det(A) \neq 0$; if A is invertible, then $\det(A^{-1}) = \frac{1}{\det(A)}$;

Properties of Determinants

- Let $A = (a_{ij})$ 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

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

- Laplace formula (Section 2.3 of Xu Yichao's book): For $1 \leq i_1 < i_2 < \cdots < i_r \leq n$,

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

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 (Section 3.2 of Xu Yichao's book): Let A and B be $m \times n$ and $n \times m$ matrices, respectively. Then:

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

Adjoint

- Let A be a square matrix of order n. Then the adjoint of A is the $n \times n$ matrix

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

  where A_{ij} is the (i, j)-cofactor of A.
- $A \operatorname{adj}(A) = \det(A) I_n$;
- If A is invertible, then $A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$;

Cramer's Rule

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

$$x = \frac{1}{\det(A)} \begin{pmatrix} \det(A_1) \\ \det(A_2) \\ \vdots \\ \det(A_n) \end{pmatrix}.$$

Exercise (2.36)

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

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

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

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

Proof and Solution.

(a) Since A is row equivalent to $\begin{pmatrix} R \\ 0 \cdots 0 \end{pmatrix}$, there exist some elementary matrices $E_1, \ldots, E_k$, such that $A = E_k \cdots E_1 \begin{pmatrix} R \\ 0 \cdots 0 \end{pmatrix}$. Hence $AB = E_k \cdots E_1 \begin{pmatrix} RB \\ 0 \cdots 0 \end{pmatrix}$, and AB can not be row equivalent to the identity matrix, i.e. AB is singular.

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

(c) For example, let $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$, then $AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is invertible.

Exercise (2.37first)

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

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

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

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

Solution.

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

(b) True. See Theorem 2.3.10.

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

$\square$

Exercise (2.40(b))

Let $A = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & -1 \\ -2 & 1 & 0 & -2 \\ 0 & 0 & 2 & 1 \end{pmatrix}$, $C = \begin{pmatrix} -1 & 3 & 4 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 2 \\ 4 \\ 6 \\ 8 \end{pmatrix}$, $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$.
Without computing the matrix AC, *explain why the homogeneous linear system* $ACx = 0$ *has infinitely many solutions.*

Solution.

- The determinant of C is the product of its diagonal entries, which is zero.
- Since $\det(AC) = \det(A)\det(C) = 0$, the homogeneous system $ACx = 0$ has infinitely many solutions.

$\square$

Exercise (2.46)

Let $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ *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* $\det(A)$*.*

Solution.

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

- If all $a, b, c, d, e, f, g, h, i$ are 1, then $\det(A) = 0$.
- Suppose at least one of $a, b, c, d, e, f, g, h, i$ is 0, say $a = 0$ (other cases are similar). Then $\det(A) = bfg + cdh - bdi - ceg$. As b, c, d, e, f, g, h, i can only be 0 and 1, $-2 \leq \det(A) \leq 2$.
- 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 $\det(A)$ is 2 and the minimum is -2.

$\square$

Exercise (2.48)

Let A be an $n \times n$ invertible matrix.

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

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

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

Proof and solution.

(a) Since $A \operatorname{adj}(A) = \det(A) I_n$ and $\det(A) \neq 0$, we have that $\operatorname{adj}(A)$ is invertible.

(b) Since $A \operatorname{adj}(A) = \det(A) I_n$, we have

$$\det(A) \det(\operatorname{adj}(A)) = \det(A \operatorname{adj}(A)) = \det(\det(A) I_n) = \det(A)^n.$$

Hence we have $\det(\operatorname{adj}(A)) = \det(A)^{n-1}$. Also by $A \operatorname{adj}(A) = \det(A) I_n$, we have that $\operatorname{adj}(A)^{-1} = \frac{1}{\det(A)} A$.

(c) First we have

$$A \cdot \operatorname{adj}(A) = \det(A) I_n, \quad \operatorname{adj}(A) \cdot \operatorname{adj}(\operatorname{adj}(A)) = \det(\operatorname{adj}(A)) I_n.$$

By part (b), we know that $\det(\operatorname{adj}(A)) = 1$. By definition, both A and $\operatorname{adj}(\operatorname{adj}(A))$ are the inverse of A, hence they are the same.

$\square$

Exercise (2.49)

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

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

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

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

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

Solution.

(a) False. For example, let $A = I_2$ and $B = -I_2$.

(b) True. See Theorem 2.5.27.

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

(d) False. For example, let $A = I_2$ and $B = C = -I_2$.

$\square$

Schedule of Today

- Any question about last tutorial
- Review concepts
- Tutorial: 3.3, 3.4, 3.7, 3.11, 3.12, 3.14
- Additional material

- n-vector: $(u_1, u_2, \ldots, u_i, \ldots, u_n)$, where u_1, u_2, …, u_n are real numbers, and u_i is the i-th coordinate.
- The set of all n-vectors of real numbers space is called the Euclidean n-space and is denoted by $\mathbb{R}^n$.
- 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: $\{n$-tuples 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: $\begin{cases} \text{Implicit form: } \{(x, y, z) \mid \text{eqn of the line}\} \\ \text{Explicit form: } \{(\text{general solution}) \mid 1 \text{ parameter}\} \end{cases}$

- $u_1, u_2, \ldots, u_k$ are fixed vectors in $\mathbb{R}^n$, and $c_1, c_2, \ldots, c_k$ are real numbers. $c_1 u_1 + c_2 u_2 + \cdots + c_k u_k$ is called a linear combination of $u_1, u_2, \ldots, u_k$.
- $S = \{u_1, u_2, \ldots, u_k\}$: a (finite) subset of $\mathbb{R}^n$. The set of all linear combinations of $u_1, u_2, \ldots, u_k$

$$\{c_1 u_1 + c_2 u_2 + \cdots + c_k u_k | c_1, c_2, \ldots, c_k \in \mathbb{R}\}$$

  is called the linear span of $u_1, u_2, \ldots, u_k$, or the linear span of S. Natation: $\text{span}\{u_1, u_2, \ldots, u_k\}$ or $\text{span}(S)$.
- Let V be a subset of $\mathbb{R}^n$. V is called a subspace of $\mathbb{R}^n$ provided there is a set $S = \{u_1, u_2, \ldots, u_k\}$ of $\mathbb{R}^n$ such that $V = \text{span}(S)$. V has a "basis" $u_1, u_2, \ldots, u_k$.
-
  - If V is a subspace of $\mathbb{R}^n$, then the zero vector $0 \in V$.
  - Let V be a subspace of $\mathbb{R}^n$. If $u_1, u_2, \ldots, u_k \in V$, and $c_1, c_2, \ldots, c_k \in \mathbb{R}$, then $c_1 u_1 + c_2 u_2 + \cdots + c_k u_k \in V$.
- By Axiom of Choice, we can prove that every "linear space" has a "basis", i.e. if a subset has the 2 properties above, then it is a subspace. Ref Exercise 3.21.
- The solution set of every homogeneous linear system is a subspace of $\mathbb{R}^n$; while the solution set of every inhomogeneous linear system is not a subspace of $\mathbb{R}^n$.

Methods for proving subspace

We have 3 methods for showing a subset to be a subspace:

- Definition: If we can find the span set, then we have done;
- As we know, the 2 necessary conditions are also sufficient, thus if we have that the subset satisfies that 2 conditions, then we have done;
- If we know the subset is a solution set of some homogeneous linear system, we have done, too;
- First 2 methods are general, while the last one is special.

Exercise (3.3)

Express each of the following by the set notation in both implicit and explicit form:

(a) *the line in* $\mathbb{R}^2$ *passing through the points* $(1, 2)$ *and* $(2, -1)$.

(b) *the plane in* $\mathbb{R}^3$ *containing the points* $(0, 1, -2)$, $(1, -1, 0)$ *and* $(0, 2, 0)$.

(c) *the line in* $\mathbb{R}^3$ *passing through the points* $(0, 1, -1)$ *and* $(1, -1, 0)$.

Recall

- To get implicit form: solving some linear system.
- To get explicit form: using implicit form, or other methods.

Solution.

(a) Substituting $(x, y) = (1, 2)$ and $(2, -1)$ into the equation $ax + by = c$, we has a system of linear equations

$$\begin{cases} a + 2b - c = 0 \\ 2a - b - c = 0 \end{cases}$$

which implies $a = \frac{3}{5}c$ and $b = \frac{1}{5}c$. In set notation, the line is $\{(x, y) \mid 3x + y = 5\}$ (implicit) and $\{(\frac{5-t}{3}, t) \mid t \in \mathbb{R}\}$ (explicit).

(b) By similar method of part (a), we have $\{(x, y, z) \mid 3x + y - z = 2\}$ (implicit) and $\{(\frac{2-s+t}{3}, s, t) \mid t \in \mathbb{R}\}$ (explicit).

(c) In explicit form, the line is
$\{(1, -1, 0) + t(-1, 2, -1) | t \in \mathbb{R}\} = \{(1 - t, -1 + 2t, -t) | t \in \mathbb{R}\}$.
To find the implicit form, we need to find two non-parallel planes containing the two points $(0, 1, -1)$ and $(1, -1, 0)$. The intersection of the two planes will give us the required line. Substituting $(0, 1, -1)$ and $(1, -1, 0)$ into $ax + by + cz = d$ we has a system of linear equations

$$\begin{cases} b - c - d = 0 \\ a - b - d = 0 \end{cases}$$

We obtain $a = c + 2d$ and $b = c + d$. There are infinitely many such planes, for example: we can write the line implicitly as $\{(x, y, z) | x + y + z = 0, 2x + y = 1\}$.

Exercise (3.4)

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

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

Which of these subsets are the same?

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, E are different.

$\square$

Exercise (3.7)

Determine which of the following are subspaces of $\mathbb{R}^4$. Justify your answers.

(a) $\{(w, x, y, z) | w + x = y + z\}$.

(b) $\{(w, x, y, z) | wx = yz\}$.

(c) $\{(w, x, y, z) | w + x + y = z^2\}$.

(d) $\{(w, x, y, z) | w = 0 \text{ and } y = 0\}$.

(e) $\{(w, x, y, z) | w = 0 \text{ or } y = 0\}$.

(f) $\{(w, x, y, z) | w = 1 \text{ and } y = 0\}$.

(g) $\{(w, x, y, z) | w = x \text{ and } y = z\}$.

(h) $\{(w, x, y, z) | w = x \text{ or } y = z\}$.

(i) $\{(w, x, y, z) | w + z = 0 \text{ and } x + y - 4z = 0 \text{ and } 4w + y - z = 0\}$.

(j) $\{(w, x, y, z) | w + z = 0 \text{ or } x + y - 4z = 0 \text{ or } 4w + y - z = 0\}$.

Recall

How to determine a subset to be a subspace? Check whether the subset satisfies the 2 properties (zero and closure).

Solution.

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

(b) No. $w_1x_1 = y_1z_1, w_2x_2 = y_2z_2 \not\Rightarrow (w_1 + w_2)(x_1 + x_2) = (y_1 + y_2)(z_1 + z_2)$, e.g. $(1, 0, 0, 1)$ and $(0, 1, 0, 1)$.

(c) No. $w_1 + x_1 + y_1 + z_1 = z_1^2, w_2 + x_2 + y_2 + z_2 = z_2^2 \not\Rightarrow$ $(w_1 + w_2) + (x_1 + x_2) + (y_1 + y_2) + (z_1 + z_2) = (z_1 + z_2)^2$, e.g. $(1, 1, -1, -1)$ and $(0, 4, 0, 2)$.

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

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

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

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

(h) No. $(1, 1, 1, 2)$ and $(1, 2, 1, 1)$ belong to the set but $(1, 1, 1, 2) + (1, 2, 1, 1) = (2, 3, 2, 3)$ does not.

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

(j) No. $(1, 0, 0, -1)$ and $(0, 0, 4, 1)$ belong to the set but $(1, 0, 0, -1) + (0, 0, 4, 1) = (1, 0, 4, 0)$ does not.

$\square$

Exercise (3.11)

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

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

(b) *Write down the subspace V explicitly if*

$$(i) A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \end{pmatrix}, \quad (ii) A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}.$$

Proof and Solution.

(a) Let $A = (c_1 \cdots c_n)$ where $c_1, \ldots, c_n$ are columns of A. For any $u = (u_1, \ldots, u_n)^T \in \mathbb{R}^n$, $Au = u_1 c_1 + \cdots + u_n c_n$. Thus $V = \mathrm{span}\{c_1, \ldots, c_n\}$ is a subspace of R^m.

(b) (i) $V = \mathbb{R}^2$. (ii) $V = \{s(1,2,3)^T + t(0,1,1)^T \mid s, t \in \mathbb{R}\}$.

$\square$

Exercise (3.12)

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

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

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

Proof and Solution.

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

(b) $A - I = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}$. 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 $\mathbb{R}^3$.

□

Exercise (3.14)

Let $V = \{(x, y, z) | x - y - z = 0\}$ *be a subset of* $\mathbb{R}^3$.

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

(b) *Let* $S = \{(1,1,0),(5,2,3)\}$. *Show that* $\mathrm{span}(S) = V$.

(c) *Let* $S' = \{(1,1,0),(5,2,3),(0,0,1)\}$. *Show that* $\mathrm{span}(S') = \mathbb{R}^3$.

Proof.

(a) Since V is a solution set of the homogeneous system, it is a subspace of $\mathbb{R}^3$. Moreover, we have that a general solution of $x - y - z = 0$ is $x = s + t$, $y = s$, $z = t$, where $s, t \in \mathbb{R}$.

□

Proof.

(b)

- Since $(1,1,0)$ and $(5,2,3)$ satisfy the equation $x-y-z=0$, $(1,1,0),(5,2,3)\in V$ and hence $\mathrm{span}(S)\subset V$.
- Let $(s+t,s,t)$ be any vector in V. We want to verify whether $a(1,1,0)+b(5,2,3)=(s+t,s,t)$ in terms of a, b can be solved.
- $a(1,1,0)+b(5,2,3)=(s+t,s,t)\Leftrightarrow\begin{cases}a+5b=s+t\\a+2b=s\\3b=t\end{cases}$
- $\left(\begin{array}{cc|c}1&5&s+t\\1&2&s\\0&3&t\end{array}\right)\xrightarrow{\text{Gaussian Elimination}}\left(\begin{array}{cc|c}1&5&s+t\\0&3&t\\0&0&0\end{array}\right)$. The system is consistent for all $s,t\in\mathbb{R}$. So $V\subset\mathrm{span}(S)$.
- Therefore $V=\mathrm{span}(S)$.

$\square$

Proof.

(c)

- It is obvious that $\text{span}(S') \subset \mathbb{R}^3$.
- Let (x, y, z) be any vector in $\mathbb{R}^3$. We want to verify whether $a(1,1,0) + b(5,2,3) + c(0,0,1) = (x, y, z)$, in terms of a, b, c, can be solved.
- $a(1,1,0) + b(5,2,3) + c(0,0,1) = (x, y, z) \Leftrightarrow \begin{cases} a + 5b = x \\ a + 2b = y \\ 3b + c = z \end{cases}$
- $\left(\begin{array}{ccc|c} 1 & 5 & 0 & x \\ 1 & 2 & 0 & y \\ 0 & 3 & 1 & z \end{array}\right) \xrightarrow{\text{Gaussian Elimination}} \left(\begin{array}{ccc|c} 1 & 5 & 0 & x \\ 0 & -3 & 0 & y - x \\ 0 & 0 & 1 & z + y - x \end{array}\right)$. The system is consistent for all $x, y, z \in \mathbb{R}$. So $\mathbb{R}^3 \subset \text{span}(S')$.
- Therefore $\mathbb{R}^3 = \text{span}(S')$.

$\square$

2nd method for part (c).

(c)
- It is obvious that $\mathrm{span}(S') \subset \mathbb{R}^3$;
- We know that $\mathbb{R}^3 = \mathrm{span}\{(1,0,0),(0,1,0),(0,0,1)\}$; It suffices to show that $(1,0,0),(0,1,0),(0,0,1)$ are linear combination of S' (It is easy);
- Therefore $\mathbb{R}^3 = \mathrm{span}(S')$.

$\square$

Page 4 Add “Ref Exercise 3.21” for 5th item;

Page 5 Add a slide for “Methods for proving subspace”;

Page 9 Change “$w + x + y + z = z^2$” to “$w + x + y = z^2$”;

Page 16 Add another method for Exercise 3.14 part(c).

Last modified: 18:41, February 23, 2010.

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 Today

- Any question about last tutorial
- Review concepts
- Tutorial: 3.18, 3.20, 3.23, 3.24(adef), 3.28, 3.37(a)
- Additional material: 3.10(a), 3.22, 3.24(g), 3.35, 3.36

Linearly independent

- Let $S = \{u_1, u_2, \dots, u_k\} \subset \mathbb{R}^n$. S is called a linearly independent set and $u_1, u_2, \dots, u_k$ are said to be linearly independent if the equation

$$c_1 u_1 + c_2 u_2 + \cdots + c_k u_k = 0$$

  has only trivial solution, where $c_1, c_2, \dots, c_k$ are variables. Otherwise, S is called a linearly dependent set and $u_1, u_2, \dots, u_k$ are said to be linearly dependent, i.e. there exist real numbers $a_1, a_2, \dots, a_k$, not all of them are zero, such that $a_1 u_1 + a_2 u_2 + \cdots + a_k u_k = 0$.
- Let $0 \in S \subset \mathbb{R}^n$, then S is linearly dependent;
- Let $u_1, u_2, \dots, u_k$ be linearly independent vectors in $\mathbb{R}^n$. Suppose u_{k+1} is a vector in $\mathbb{R}^n$, and not a linear combination of $u_1, u_2, \dots, u_k$. Then $u_1, u_2, \dots, u_k, u_{k+1}$ are linearly independent.

Linearly independent

How to determine whether a subset is linearly independent?

- 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 $S = \{u\} \subset \mathbb{R}^n$, then S is linearly dependent iff $u = 0$;
- Let $S = \{u, v\} \subset \mathbb{R}^n$, then S is linearly dependent iff $u = av$ for some $a \in \mathbb{R}$ or $v = bu$ for some $b \in \mathbb{R}$;
- Let $S = \{u_1, u_2, \dots, u_k\} \subset \mathbb{R}^n$ where $k \geq 2$, then
  - S is linearly dependent iff at least one vector u_i in S can be written as a linear combination of the other vectors in S;
  - S is linearly independent iff no vector u_i in S can be written as a linear combination of the other vectors in S.
- Let $S = \{u_1, u_2, \dots, u_k\} \subset \mathbb{R}^n$. If $k > n$, then S is linearly dependent.
- In $\mathbb{R}^n$, 2 vectors u, v are linearly dependent iff they lie on the same line.
- In $\mathbb{R}^n$, 3 vectors u, v, w are linearly dependent iff they lie on the same plane.

Bases

- Let $S = \{u_1, u_2, \ldots, u_k\}$ be a subset of a vector space V. Then S is called a basis for V if
  - S is linearly independent;
  - S spans V.
- Let $S = \{u_1, u_2, \ldots, u_k\}$ be a basis for a vector space V, then every vector $v \in V$ can be expressed in the form $v = c_1 u_1 + c_2 u_2 + \cdots + c_k u_k$ in exactly one way, where $c_1, c_2, \ldots, c_k \in \mathbb{R}$.
- Let $S = \{u_1, u_2, \ldots, u_k\}$ be a basis for a vector space V and $v \in V$. If $v = c_1 u_1 + c_2 u_2 + \cdots + c_k u_k$, then the coefficients $c_1, c_2, \ldots, c_k$ are called the coordinates of v relative to the basis S. The vector $(v)_S = (c_1, c_2, \ldots, c_k) \in \mathbb{R}^k$ is called the coordinate vector of v relative to the basis S.
- $S = \{e_1, e_2, \ldots, e_n\}$, where $e_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^n$, is the standard basis for $\mathbb{R}^n$, and $(u)_S = (u_1, u_2, \ldots, u_n) = u$.
-
  - Suppose $S \subset V$ and $\mathrm{span}(S) = V$, then there exists $S' \subset S$, such that S' is a basis for V;
  - Suppose T is a set of linearly independent vectors in V. Then there exists a basis T' for V such that $T \subset T'$.

Dimension

- 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.
- 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.

Exercise (3.18)

Let u, v, w *be the vectors and let*

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

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

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

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

Solution of (a).

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

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

  Hence S_3 spans $\mathbb{R}^3$.
- Using the same argument as for S_3, we can show that both S_4 and S_5 also span $\mathbb{R}^3$.

□

Solution of (b).

- If there exist $a, b \in \mathbb{R}$, which are not both 0, such that $au + bv = 0$, then $au + bv + 0w = 0$, i.e. u, v, w are linearly dependent, contradiction.
- Since $(u - v) + (v - w) + (w - u) = 0$, they are linearly dependent.
- Suppose $a(u - v) + b(v - w) + c(w + u) = 0$, it is equivalent to $(a + c)u + (-a + b)v + (-b + c)w = 0$. Since u, v, 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. $u - v, v - w, w + u$ are linearly independent.
- By similarly method, we have that S_4 and S_5 are linearly independent.

$\square$

Exercise (3.20)

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

(a) u, v, w *are linearly independent.*

(b) u, v, w *are not linearly independent.*

Solution.

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

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

$\square$

Exercise (3.23)

(All vectors in this question are column vectors.) Let $u_1, u_2, \ldots, u_k$ be vectors in $\mathbb{R}^n$ and A an $n \times n$ matrix.

(a) *Show that if $Au_1, Au_2, \ldots, Au_k$ are linearly independent, then $u_1, u_2, \ldots, u_k$ are linearly independent.*

(b) *Suppose $u_1, u_2, \ldots, u_k$ are linearly independent.*
- *Show that if A is invertible, then $Au_1, Au_2, \ldots, Au_k$ are linearly independent.*
- *If A is not invertible, are $Au_1, Au_2, \ldots, Au_k$ linearly independent?*

Proof of (a).

Suppose $c_1 u_1 + c_2 u_2 + \cdots + c_k u_k = 0$, then

$$c_1 A u_1 + c_2 A u_2 + \cdots + c_k A u_k = A(c_1 u_1 + c_2 u_2 + \cdots + c_k u_k) = 0.$$

Hence $c_1 = c_2 = \cdots = c_k = 0$, i.e. $u_1, u_2, \ldots, u_k$ are linearly independent. $\square$

Proof of (b).

- - Suppose $c_1Au_1 + c_2Au_2 + \cdots + c_kAu_k = 0$, then $A(c_1u_1 + c_2u_2 + \cdots + c_ku_k) = 0$.
  - Since A is invertible, $c_1u_1 + c_2u_2 + \cdots + c_ku_k = 0$.
  - Since $u_1, u_2, \ldots, u_k$ are linearly independent, we have $c_1 = c_2 = \cdots = c_k = 0$, i.e. $Au_1, Au_2, \ldots, Au_k$ are linearly independent.
- No conclusion. For example,
  - let $u_1 = (1, 0, 0)^T$ and $u_2 = (0, 1, 0)^T$: It is obvious that u_1 and u_2 are linearly independent.
  - If $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, then Au_1 and Au_2 are linearly independent.
  - If $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, then Au_1 and Au_2 are linearly dependent.

$\square$

Exercise (3.24(adef))

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

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

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

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

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

Proof.

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

(d) False. For example, let $u = (1, 1), v = (2, 2)$.

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

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

$\square$

Exercise (3.28)

Let $V = \{(a+b, a+c, c+d, b+d) | a, b, c, d \in \mathbb{R}\}$ *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* $u = (1,2,3,2)$ *relative to* S.

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

Proof.

(a) $V = \{a(1,1,0,0) + b(1,0,0,1) + c(0,1,1,0) + d(0,0,1,1) | a, b, c, d \in \mathbb{R}\} = \mathrm{span}\{(1,1,0,0), (1,0,0,1), (0,1,1,0), (0,0,1,1)\}$ and hence is a sub-space of $\mathbb{R}^4$. It is obvious that S is linearly independent and $\mathrm{span}(S) = \mathrm{span}\{(1,1,0,0), (1,0,0,1), (0,1,1,0), (0,0,1,1)\} = V$. 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 $c_1 = 4, c_2 = -3, c_3 = 2$, i.e. the coordinate vector of u relative to S is $(4,-3,2)$.

(c) $v = 1(1,1,0,0) + 3(1,0,-1,0) - 1(0,-1,0,1) = (4,2,-3,-1)$.

□

Exercise (3.37(a))

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 $V \cap W$.*

Proof.

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

$\square$

Exercise (3.10(a))

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

Proof.

It is obvious that $V + W$ satisfies these conditions, i.e. it is a subspace of $\mathbb{R}^n$. □

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 and Solution.

(a) It is obvious that $V \cap W$ satisfies these conditions, i.e. it is a subspace of $\mathbb{R}^n$.

(b) Let $V = \{(x, 0) | x \in \mathbb{R}\}$ and $W = \{(0, y) | y \in \mathbb{R}\}$. 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 $\mathbb{R}^n$.

(c) Suppose $V \not\subset W$. We want to show that $W \subset V$. Take any vector $x \in W$, we want to show $x \in V$. Since $V \not\subset W$, there exists a vector $y \in V$ but $y \notin W$. Then since $V \cup W$ is a subspace of $\mathbb{R}^n$ and $x, y \in V \cup W$, we have $x + y \in V \cup W$, i.e. either $x + y \in V$ or $x + y \in W$.

- Assume $x + y \in W$. As W is a subspace of $\mathbb{R}^n$, we have $y = (x + y) - x \in W$ which contradict that $y \notin W$ as mentioned above.
- Now we know that $x + y \in V$. As V is a subspace of $\mathbb{R}^n$, we have $x = (x + y) - y \in V$.

Since every vector in W must be contained in V, $W \subset V$.

□

Exercise (3.24(g))

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

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

Solution.

True.

- For any element u of $\mathrm{span}(S_1 \cup S_2)$, it can be expressed as a linear combination of $S_1 \cup S_2$. Hence, $u = u_1 + u_2$ where $u_1 \in \mathrm{span}(S_1)$ and $u_2 \in \mathrm{span}(S_2)$.
- For any elements $u_1 \in \mathrm{span}(S_1)$ and $u_2 \in \mathrm{span}(S_2)$, $u_1 + u_2$ is a linear combination of $S_1 \cup S_2$.

$\square$

Exercise (3.35)

Let V be a vector space of dimension of n. Show that there exists $n+1$ vectors $u_1, u_2, \dots, u_n, u_{n+1}$ such that every vector in V can be expressed as a linear combination of $u_1, u_2, \dots, u_{n+1}$ with non-negative coefficients.

Proof.

- Take a basis $\{u_1, u_2, \dots, u_n\}$ for V. Define $u_{n+1} = -u_1 - u_2 - \cdots - u_n$.
- For any $v \in V$, $v = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n$ for some $a_1, a_2, \dots, a_n \in \mathbb{R}$.
- Let $a = \min\{0, a_1, a_2, \dots, a_n\}$. Then $v = (a_1 - a)u_1 + (a_2 - a)u_2 + \cdots + (a_n - a)u_n + (-a)u_{n+1}$ where $a_i - a \geq 0$, for $i = 1, 2, \dots, n$, and $-a \geq 0$.
- So every vector in V can be expressed as a linear combination of $u_1, u_2, \dots, u_n, u_{n+1}$ with non-negative coefficients.

$\square$

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 $\{u_1, \ldots, u_k\}$ be a basis for $V \cap W$. By Problem 3.4.8.2, there exists vectors $v_1, \ldots, v_m \in V$ such that $\{u_1, \ldots, u_k, v_1, \ldots, v_m\}$ is a basis for V and there exists vectors $w_1, \ldots, w_n \in W$ such that $\{u_1, \ldots, u_k, w_1, \ldots, w_n\}$ is a basis for W. It is easy to see that $V + W = \operatorname{span}\{u_1, \ldots, u_k, v_1, \ldots, v_m, w_1, \ldots, w_n\}$.
- Consider $a_1u_1 + \cdots + a_ku_k + b_1v_1 + \cdots + b_mv_m + c_1w_1 + \cdots + c_nw_n = 0 (*)$. Since $c_1w_1 + \cdots + c_nw_n = -(a_1u_1 + \cdots + a_ku_k + b_1v_1 + \cdots + b_mv_m) \in V \cap W$, there exists $d_1, \ldots, d_k \in \mathbb{R}$ such that $c_1w_1 + \cdots + c_nw_n = d_1u_1 + \cdots + d_ku_k$, i.e. $c_1w_1 + \cdots + c_nw_n - d_1u_1 - \cdots - d_ku_k = 0$. 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 $a_1u_1 + \cdots + a_ku_k + b_1v_1 + \cdots + b_mv_m = 0$. As $\{u_1, \ldots, u_k, v_1, \ldots, v_m\}$ is linearly independent, $a_1 = \cdots = a_k = b_1 = \cdots = b_m = 0$.
- So $(*)$ has only the trivial solution and hence $\{u_1, \ldots, u_k, v_1, \ldots, v_m, w_1, \ldots, w_n\}$ is linearly independent. We have shown that $\{u_1, \ldots, u_k, v_1, \ldots, v_m, w_1, \ldots, w_n\}$ is a basis for $V + W$.
- Thus $\dim(V + W) = k + m + n = \dim(V) + \dim(W) - \dim(V \cap W)$.

□

Exercise (3.37(bcd))

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

(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.

(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.

$\square$

Page 4 Change "Let $S = \{u, v\} \subset \mathbb{R}^2$" to "Let $S = \{u, v\} \subset \mathbb{R}^n$";

Page 5 Change "$T \subset V'$" to "$T \subset T'$";

Page 11 Change "$Au_1, Au_2, \ldots, Au_k$ are linearly independent" in the Proof of (a) to "$u_1, u_2, \ldots, u_k$ are linearly independent";

Last modified: 09:15, March 1st, 2010.

Schedule of Today

- Any question about last tutorial
- Review concepts
- Tutorial: 3.31, 3.33, 3.34, 4.7, 4.8, 4.18
- Additional material

Dimension

Thm 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.

Def 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.

Thm 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 $|S| = k = \dim(V)$.

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

(1) A is invertible;
(2) The linear system $Ax = 0$ has only trivial solution;
(3) The RREF of A is an identity matrix;
(4) A can be expressed as a product of elementary matrices;
(5) $\det(A) \neq 0$;
(6) The rows of A form a basis for $\mathbb{R}^n$;
(7) The columns of A form a basis for $\mathbb{R}^n$.

Transition matrices

- Let $S = \{u_1, u_2, \ldots, u_k\}$ be a basis for a vector space V and v be a vector in V. If $v = c_1 u_1 + c_2 u_2 + \cdots + c_k u_k$, then the vectors

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

  are called the coordinate vector of v relative to S.
- Let $S = \{u_1, u_2, \ldots, u_k\}$ and $T = \{v_1, v_2, \ldots, v_k\}$ be two bases for a vector space V. Then matrix

$$P = \big([u_1]_T, [u_2]_T, \ldots, [u_k]_T\big)$$

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

Row spaces and Column spaces

- Let A be an $m \times n$ matrix. The row space of A is the subspace of $\mathbb{R}^n$ spanned by the rows of A. The column space of A is the subspace of $\mathbb{R}^m$ spanned by the columns of A.
- Let A and B be row equivalent matrices, then the row space of $A =$ the row space of B.
- Let A and B be row equivalent matrices. Then the following statements hold:
  - A given set of columns of A is linearly independent iff the set of corresponding columns of B is linearly independent;
  - A given set of columns of A forms a basis for the column space of A iff the set of corresponding columns of B forms a basis for the column space of B.

Exercise (3.31)

Let $\{u_1, u_2, u_3\}$ be a basis for a vector space V. Determine whether $\{v_1, v_2, v_3\}$ is a basis for V if

(a) $v_1 = u_1$, $v_2 = u_1 + u_2$, $v_3 = u_1 + u_2 + u_3$.

(b) $v_1 = u_1 - u_2$, $v_2 = u_2 - u_3$, $v_3 = u_3 - u_1$.

Solution.

(a)
- Suppose $c_1 v_1 + c_2 v_2 + c_3 v_3 = 0$. Then
$$(c_1 + c_2 + c_3)u_1 + (c_2 + c_3)u_2 + c_3 u_3 = 0.$$
- Since u_1, u_2, u_3 are linearly independent, $c_1 + c_2 + c_3 = c_2 + c_3 = c_3 = 0$, i.e. $c_1 = c_2 = c_3 = 0$.
- Hence, v_1, v_2, v_3 are linearly independent.
- Since $\dim(V) = 3$, $\{v_1, v_2, v_3\}$ is a basis for V.

(b) Since $v_1 + v_2 + v_3 = 0$, they are linearly dependent. Hence, $\{v_1, v_2, v_3\}$ is not a basis for V.

$\square$

Exercise (3.33)

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

(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.

(a)
- Since $V = \{(x, y, z) | 2x - y + z = 0\} = \{(x, 2x + z, z) | x, z \in \mathbb{R}\} = \text{span}\{(1, 2, 0), (0, 1, 1)\} = \text{span}(S)$, 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 $\dim(V) = |S| = 2$, 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 (b).

$$\left(\begin{array}{cc|c|c} 0 & 1 & 1 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 0 & -1 & -2 \end{array}\right) \xrightarrow{\text{GJ Elimination}} \left(\begin{array}{cc|c|c} 1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$ Thus $[(1,1,-1)]_S = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $[(1,0,-2)]_S = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$. The transition matrix from T to S is $\begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}$. The transition matrix from S to T is

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

□

Exercise (3.34)

Let $S = \{u_1, u_2, u_3\}$ *be a basis for* $\mathbb{R}^3$ *and* $T = \{v_1, v_2, v_3\}$ *where* $v_1 = u_1 + u_2 + u_3, v_2 = u_2 + u_3, v_3 = u_2 - u_3$.

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

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

Proof and Solution.

(a)
- Suppose $c_1 v_1 + c_2 v_2 + c_3 v_3 = 0$, then $c_1 u_1 + (c_1 + c_2 + c_3) u_2 + (c_1 + c_2 - c_3) u_3 = 0$.
- Since u_1, u_2, u_3 are linearly independent, we have $c_1 = c_1 + c_2 + c_3 = c_1 + c_2 - c_3 = 0$, i.e. $c_1 = c_2 = c_3 = 0$. Hence, v_1, v_2, v_3 are linearly independent.
- Since $\dim(\mathbb{R}^3) = 3 = |T|$, T is a basis for V.

(b) We know that $[v_1]_S = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $[v_2]_S = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, $[v_3]_S = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$. The transition matrix from T to S is $P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$, and the transition matrix from S to T is $P^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1/2 & 1/2 \\ 0 & 1/2 & -1/2 \end{pmatrix}$.

Exercise (4.7)

Let $V = \mathrm{span}\{u_1, u_2, u_3, u_4\}$ *where*

$$u_1 = (1,1,1,1,1), u_2 = (1,x,x,x,x), u_3 = (1,x,x^2,x,x^2), u_4 = (1,x^3,x,2x-x^3,x)$$

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

Solution.

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & x & x & x & x \\ 1 & x & x^2 & x & x^2 \\ 1 & x^3 & x & 2x-x^3 & x \end{pmatrix} \xrightarrow{\text{GE}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & x-1 & x-1 & x-1 & x-1 \\ 0 & 0 & x^2-x & 0 & x^2-x \\ 0 & 0 & 0 & 2x-2x^3 & 0 \end{pmatrix}$$

- If $x = 1$, then $\{u_1\}$ is a basis for V and $\dim(V) = 1$.
- If $x = 0$, then $\{u_1, (0,1,1,1,1)\}$ is a basis for V and $\dim(V) = 2$.
- If $x = -1$, then $\{u_1, (0,-2,-2,-2,-2), (0,0,2,0,2)\}$ is a basis for V and $\dim(V) = 3$.
- If $x \notin \{0,1,-1\}$, then $\{u_1, u_2, u_3, u_4\}$ is a basis for V and $\dim(V) = 4$.

$\square$

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 (4.18)

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* A*.*

(b) *Find a basis for the row space of* A *and a basis for the column space of* A*.*

Solution.

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

(b) It is obvious that $\{a_1, a_2, a_3\}$ is a basis for the column space of 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 A.

□

Schedule of Today

- Any question about last tutorial
- Review concepts: Rank, Nullity
- Tutorial: 4.13, 4.16, 4.17, 4.20, 4.21, 4.23
- Additional material: 4.22, 4.24, 4.25, 4.26, 4.27, 7a, 7b, 7c, 7d

Rank and Nullity

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

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

- If R is a REF of A, then

$$\begin{aligned}\mathrm{rank}(A) &= \#\text{ non-zero rows of } R = \#\text{ leading entries of } R = \#\text{ pivot columns of } R \\ &= \text{largest } \#\text{ of L.I. rows in } A = \text{largest } \#\text{ of L.I. columns in } A \\ &= \text{largest size of invertible submatrices of } A \text{ (See later)}\end{aligned}$$

- $A \in \mathbb{R}^{m\times n}$, then $\mathrm{rank}(A) \le \min\{m, n\}$.
- $A \in \mathbb{R}^{m\times n}$, then $\mathrm{rank}(A) = \mathrm{rank}(A^T)$.

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

- $A \in \mathbb{R}^{m\times n}$, then $\mathrm{rank}(A) + \mathrm{nullity}(A) = (\#\text{ columns of } A) = n$.
- B is a submatrix of A, then $\mathrm{rank}(B) \le \mathrm{rank}(A)$.

(1) $\mathrm{rank}(AB) \le \min\{\mathrm{rank}(A), \mathrm{rank}(B)\}$. See Exercise 4.23.

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

(2a) $A \in \mathbb{R}^{m\times n}$, $\mathrm{rank}(A) = r \le \min\{m, n\}$, then there exist invertible matrices $P \in \mathbb{R}^{m\times m}$ and $Q \in \mathbb{R}^{n\times n}$, such that $PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. By (2).

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

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

(3) $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{p\times q}$, $C \in \mathbb{R}^{m\times p}$, then $\mathrm{rank}\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \ge \mathrm{rank}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$. Def.

(3a) $A \in \mathbb{R}^{m\times p}$, $B \in \mathbb{R}^{p\times n}$, then $\mathrm{rank}(AB) \ge \mathrm{rank}(A) + \mathrm{rank}(B) - p$. By (2), (3).

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

(4) $\mathrm{rank}(A \pm B) \le \mathrm{rank}(A) + \mathrm{rank}(B)$. By (2), (2b) and Def.

(4a) $\mathrm{rank}\,(A, B) \le \mathrm{rank}(A) + \mathrm{rank}(B)$, $\mathrm{rank}\begin{pmatrix} A \\ B \end{pmatrix} \le \mathrm{rank}(A) + \mathrm{rank}(B)$. By (4).

(4b) $\mathrm{rank}(A - B) \ge |\,\mathrm{rank}(A) - \mathrm{rank}(B)|$. By (4).

(4c) $A \in \mathbb{R}^{n\times n}$, then $\mathrm{rank}(A) + \mathrm{rank}(I_n + A) \ge n$. By (4).

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

Exercise (4.13)

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

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

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

Solution.

(a) By Gauss 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, i.e., $\text{rank}(A) = 1, \text{nullity}(A) = 2$;
- when $a = -2$, there are 2 non-zero rows, i.e., $\text{rank}(A) = 2, \text{nullity}(A) = 1$;
- For other cases, all of the rows are non-zero rows, i.e., $\text{rank}(A) = 3, \text{nullity}(A) = 0$.

□

Solution for (b).

For

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

- the first 2 rows are linearly independent, then $\mathrm{rank}(B) \leq 2$.
- if $b = c = d = e = f = 0$, $\mathrm{rank}(B) = 0, \mathrm{nullity}(B) = 3$;
- 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, $\mathrm{rank}(B) = 1, \mathrm{nullity}(B) = 2$.
- if not all b, c are zero and not all d, e are zero, $\mathrm{rank}(B) = 2, \mathrm{nullity}(B) = 1$.

$\square$

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 (4.16)

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

(a) *Find a* 4×4 *matrix* A *such that the row space of* A *is* V.

(b) *Find a* 4×4 *matrix* B *such that the column space of* B *is* V.

(c) *Find a* 4×4 *matrix* C *such that the nullspace of* C *is* V.

Solution.

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

(c)
- Since $(1,2,0,0), (0,-1,1,0), (0,0,0,1)$ are linearly independent, then $\dim(V) = 3$, i.e. the rank of $C = (c_{i,j})_{4\times 4}$ which need to find is 1. So we can let the last 3 rows of C be zero rows. Now it suffices to find $c_{11}, c_{12}, c_{13}, c_{14}$.
- Since $C(1,2,0,0)^T = C(0,-1,1,0)^T = C(0,0,0,1)^T = \vec{0}$, then $c_{11} + 2c_{12} = 0$, $c_{12} - c_{13} = 0$, $c_{14} = 0$. Then we can choose

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

□

Exercise (4.17)

Let 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 $Ax = b$ and that the homogeneous system $Ax = 0$ 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 A and determine the nullity of A.*

(b) *Find a general solution for the system $Ax = b$.*

(c) *Write down the RREF of A.*

(d) *Find a basis for the row space of A and determine the rank of A.*

(e) *Do we have enough information for us to find the column space of A?*

Solution.

(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 A. The nullity of A is 2.

(b) A general solution of $Ax = b$ 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 (cde).

(c)
- It is obvious that $\mathrm{nullity}(A) = 2$, and $\mathrm{rank}(A) = 1$. So we have that the last row in the RREF of A is a zero row.
- A general solution of $Ax = 0$ 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 $\mathrm{rank}(A) = 2$) 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 ($i = 1, 2$), i.e. 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 A. The rank of A is 2.

(e) No, we cannot find the column space of A with the given information.

□

Exercise (4.20)

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

Proof.

- Let $B = (\vec{b_1}, \ldots, \vec{b_n})$ where $\vec{b_j}$ is the j-th column of B.
-
$$AB = 0 \Rightarrow (A\vec{b_1}, \ldots, A\vec{b_n}) = 0 \Rightarrow A\vec{b_j} = \vec{0} \text{ for all } j,$$

  $\vec{b_1}, \ldots, \vec{b_n}$ are contained in the nullspace of A.
- So the column space of $B = \operatorname{span}\{\vec{b_1}, \ldots, \vec{b_n}\}$ the nullspace of A.

$\square$

Exercise (4.21)

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

Proof.

- Let $A = \begin{pmatrix} \vec{a_1} \\ \vdots \\ \vec{a_n} \end{pmatrix}$ be a matrix where $\vec{a_i}$ is the i-th row of A. Let $\vec{u}$ be any column vector in the nullspace of A. Then

$$A\vec{u} = \vec{0} \Rightarrow \begin{pmatrix} \vec{a_1}\,\vec{u} \\ \vdots \\ \vec{a_n}\,\vec{u} \end{pmatrix} = 0 \Rightarrow \vec{a_i}\vec{u} = 0 \text{ for all } i.$$

- Let $\vec{b}$ be any vector in the row space of A, i.e. $\vec{b} = c_1\,\vec{a_1} + \cdots + c_n\,\vec{a_n}$ where $c_1, \ldots, c_n$ are scalars. We have $\vec{b}\vec{u} = c_1\,\vec{a_1}\,\vec{u} + \cdots + c_n\,\vec{a_n}\,\vec{u} = 0$.
- 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 (4.23)

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

(a) *Show that the nullspace of B is a subset of the nullspace of AB. Hence prove that $\mathrm{rank}(AB) \leq \mathrm{rank}(B)$.*

(b) *Show that every column of the matrix AB lies in the column space of A. Hence, or otherwise, prove that $\mathrm{rank}(AB) \leq \mathrm{rank}(A)$.*

Proof.

(a) Let $\vec{u}$ be any vector in the nullspace of B, i.e. $B\vec{u} = 0$. Since $AB\vec{u} = A0 = 0$, $\vec{u}$ is also a vector in the nullspace of AB. So we have shown that the nullspace of B is a subset of the nullspace of AB. Since $\mathrm{nullity}(B) \leq \mathrm{nullity}(AB)$,

$$\mathrm{rank}(AB) = n - \mathrm{nullity}(AB) \leq n - \mathrm{nullity}(B) = \mathrm{rank}(B).$$

(b) By (a), $\mathrm{rank}(B^T A^T) \leq \mathrm{rank}(A^T)$. Since $\mathrm{rank}(A^T) = \mathrm{rank}(A)$ and $\mathrm{rank}(AB) = \mathrm{rank}(B^T A^T)$, $\mathrm{rank}(AB) \leq \mathrm{rank}(A)$.

□

Exercise (4.22)

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

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

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

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

Proof.

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

(b) $P = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, $A = I_2$, then $\operatorname{rank}(PA) = 0 \neq 2 = \operatorname{rank}(A)$.

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

$\square$

Exercise (4.24)

Let A and B be two matrices of the same size. Show that

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

Proof.

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

□

Exercise (4.25)

Let A be an $m \times n$ matrix.

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

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

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

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

(a) Proved in lecture;

(b) By (a);

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

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

Exercise (4.26)

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

Proof.

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

$\square$

Exercise (4.27)

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

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

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

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

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

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

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

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

Proof.

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

(b) True. Since the row space of A and the the row space of B are the same. Hence the column space of A^T and the column space of B^T are the same.

(d) False. For example, let $A = B = I_1$.

(e) False. For example, let $A = B = 0_1$.

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

$\square$

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

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

  (i) *Show that the nullspace of A is equal to the nullspace of A^2.*

  (ii) *Show that the nullspace of A and the column space of A intersect trivially.*

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

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

  (i) *Find the nullspace of B^2.*

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

Exercise (7a)

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

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

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

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

Exercise (7b)

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

Exercise (7c)

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

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

Exercise (7d)

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

- Page 1 Change "$\text{rank}(A) + \text{nullity}(A) = n$" to "$\text{rank}(A) + \text{nullity}(A) = $ (# columns of A) $= n$". Add 1 property "B is a submatrix of A, then $\text{rank}(B) \leq \text{rank}(A)$.";
- Page 6 Add 2 additional questions;
- Page 7 Revise the solution for part (c);
- Page 9 Revise the solution for part (c);
- Page 19 Add 1 additional question;
- Page 20 Add 1 part for Exercise (7c).

Last modified: 13:12, March 16st, 2010.

Schedule of Today

- Any question about last tutorial
- Review concepts
- Tutorial: 6.3, 6.7(c), 6.11, 6.12, 6.14(a), 6.17(b)
- Additional material: the algebraic multiplicity, the geometric multiplicity

Eigenvalue

Here we only consider real case. Let $\boldsymbol{A}$ be a 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 some invertible matrix, then $\boldsymbol{A}$ and $\boldsymbol{B}$ have same eigenvalues. While the converse is not necessarily true. (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 $\boldsymbol{A}$ has n eigenvalues $\{\lambda_i\}_{i=1}^n$, then $\mathrm{tr}(\boldsymbol{A}) = \sum_{i=1}^n \lambda_i$, $\det(\boldsymbol{A}) = \prod_{i=1}^n \lambda_i$. (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 $\varphi(\boldsymbol{A}) = \boldsymbol{0}$. (Exercise 6.2(b))

Algebraic multiplicity, Geometric multiplicity

Let $\boldsymbol{A}$ be a square matrix of order n, the characteristic polynomial be
$\varphi_{\boldsymbol{A}}(\lambda) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_k)^{r_k}(\lambda^2 + a_1\lambda + b_1)^{s_1} \cdots (\lambda^2 + a_l\lambda + b_l)^{s_l}$.

- 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 $\boldsymbol{A}$,

$$\text{the algebraic multiplicity of } \lambda \geq \text{ the geometric multiplicity of } \lambda \geq 1.$$

- A square matrix $\boldsymbol{A}$ is called diagonalizable if there exists an invertible matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is a diagonal matrix.
- Let $\boldsymbol{A}$ be a square matrix of order n. Then $\boldsymbol{A}$ is diagonalizable iff $\boldsymbol{A}$ has n linearly independent eigenvectors.
- How to determine whether a square matrix is diagonalizable?
  (1) Decompose the characteristic polynomial as $\varphi_{\boldsymbol{A}}(\lambda) = (\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_k)^{r_k}(\lambda^2+a_1\lambda+b_1)^{s_1}\cdots(\lambda^2+a_l\lambda+b_l)^{s_l}$, where $\lambda_1,\ldots,\lambda_k$ are pairwise distinct, $(\lambda^2+a_j\lambda+b_j)$ can not do more decomposition. If $k=n$, then $\boldsymbol{A}$ is diagonalizable; 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.
- Let $\boldsymbol{A}$ be an complex matrix, then there exists an invertible matrix $\boldsymbol{P}$, such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is an upper-triangle matrix.

Conditions for diagonalizability

- The square matrix $\boldsymbol{A}$ of order n is diagonalizable iff $\boldsymbol{A}$ has n linearly independent eigenvectors.
- If the square matrix $\boldsymbol{A}$ of order n 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.

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.

$$
\begin{aligned}
&\lambda \text{ is an eigenvalue of } \boldsymbol{A} \\
\Rightarrow &\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0 \\
\Rightarrow &\det\left((\lambda \boldsymbol{I} - \boldsymbol{A})^T\right) = 0 \\
\Rightarrow &\det(\lambda \boldsymbol{I} - \boldsymbol{A}^T) = 0 \\
\Rightarrow &\lambda \text{ is an eigenvalue of } \boldsymbol{A}^T
\end{aligned}
$$

□

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.

(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.

(c)
- 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.
- 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.
- In both cases, we can find a nonzero vector $\boldsymbol{u} \in E_2 \cap E'_\lambda$, i.e. $\boldsymbol{A}\boldsymbol{u} = 2\boldsymbol{u}$ and $\boldsymbol{B}\boldsymbol{u} = \lambda\boldsymbol{u}$, such that
$$(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{u} = \boldsymbol{A}\boldsymbol{u} + \boldsymbol{B}\boldsymbol{u} = 2\boldsymbol{u} + \lambda\boldsymbol{u} = (2+\lambda)\boldsymbol{u}.$$
- So $2+\lambda$ is an eigenvalue of $\boldsymbol{A}+\boldsymbol{B}$.

□

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.

□

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 for (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 $\boldsymbol{A}$. $\square$

Proof for (ii).

By Question 6.3, λ is an eigenvalue of $\boldsymbol{A}^T$. Let $x = (x_1, x_2, \ldots, x_n)^T$ be a eigenvector of $\boldsymbol{A}^T$ associated with the eigenvalue λ, i.e. $\boldsymbol{A}^T\boldsymbol{x} = \lambda\boldsymbol{x}$. Choose $k \in \{1, 2, \ldots, n\}$ such that $|x_k| = \max\{|x_i| : i = 1, 2, , \ldots, n\}$, i.e. $|x_k| \geq |x_i|$ for $i = 1, 2, \ldots, n$. Since $\boldsymbol{x}$ is a nonzero vector, $|x_k| > 0$.
By comparing the k-th coordinate of both sides of $\boldsymbol{A}^T\boldsymbol{x} = \lambda\boldsymbol{x}$, we have

$$a_{1k}x_1 + a_{2k}x_2 + \cdots + a_{nk}x_n = \lambda x_k.$$

Hence,

$$\begin{aligned} |\lambda||x_k| &= |a_{1k}x_1 + a_{2k}x_2 + \cdots + a_{nk}x_n| \\ &\leq |a_{1k}x_1| + |a_{2k}x_2| + \cdots + |a_{nk}x_n| \\ &\leq a_{1k}|x_1| + a_{2k}|x_2| + \cdots + a_{nk}|x_n| && (a_{ij} \geq 0) \\ &\leq (a_{1k} + a_{2k} + \cdots + a_{nk})|x_k| \end{aligned}$$

So $|\lambda| \leq 1$. $\square$

Exercise (6.17(b))

Following the procedure discussed in Example 6.2.9.2 or Example 6.2.12, solve the following recurrence relation $a_n = a_{n-1} + 2a_{n-2}$ *with* $a_0 = 1$ *and* $a_1 = 0$.

Solution.

- Since $\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix}$, let $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}$. It is easy to get that $\boldsymbol{A}$ has 2 distinct eigenvalues 2 and -1. Hence there exists an invertible matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$.
- Thus
$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \boldsymbol{P} \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}^n \boldsymbol{P}^{-1} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} b2^n + c(-1)^n \\ d2^n + e(-1)^n \end{pmatrix}$$
for some constants b, c, d, e.
- In fact, $a_n = b2^n + c(-1)^n$. Since $a_0 = 1$ and $a_1 = 0$, we obtain $b = \frac{1}{3}$ and $c = \frac{2}{3}$. Thus $a_n = \frac{1}{3}[2^n + 2(-1)^n]$.

□

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.

- 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 $\mathbb{R}^n$: $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m, \boldsymbol{\alpha}_{m+1}, \ldots, \boldsymbol{\alpha}_n\}$.
- $$\boldsymbol{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m, \ldots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m, \ldots, \boldsymbol{\alpha}_n) \begin{pmatrix} \lambda_0 \boldsymbol{I}_m & \boldsymbol{B} \\ 0 & \boldsymbol{C} \end{pmatrix}$$
- Then $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - \lambda_0)^m \det(\lambda \boldsymbol{I}_{n-m} - \boldsymbol{C})$.
- Hence, the algebraic multiplicity of some eigenvalue λ_0 is greater then or equal to the geometric multiplicity of λ.

□

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.

1st Method.

- 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.
- 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 $a\lambda_1\boldsymbol{x}_1 + b\lambda_1\boldsymbol{x}_2 = \mathbf{0}$.
- Then we will get $b(\lambda_1 - \lambda_2)\boldsymbol{x}_2 = \mathbf{0}$, i.e., $b = 0$. Similarly, $a = 0$. Contradiction.
- For general case, we can apply mathematical induction, left it for you.

$\square$

2nd Method.

- 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 $a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_t\boldsymbol{x}_t = \boldsymbol{0}$, where not all of $a_1, \ldots, a_t$ are zero.
- Then $\boldsymbol{0} = \boldsymbol{A} \cdot \boldsymbol{0} = \boldsymbol{A}(a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_t\boldsymbol{x}_t) = a_1\lambda_1\boldsymbol{x}_1 + a_2\lambda_2\boldsymbol{x}_2 + \cdots + a_t\lambda_t\boldsymbol{x}_t$.
- Similarly, we have $a_1\lambda_1^2\boldsymbol{x}_1 + a_2\lambda_2^2\boldsymbol{x}_2 + \cdots + a_t\lambda_t^2\boldsymbol{x}_t = \boldsymbol{0}$.
- By induction, we have $a_1\lambda_1^j\boldsymbol{x}_1 + a_2\lambda_2^j\boldsymbol{x}_2 + \cdots + a_t\lambda_t^j\boldsymbol{x}_t = \boldsymbol{0}$ for $j = 1, 2, \ldots, t$.
- Consider the linear system: $\begin{cases} a_1\lambda_1 y_1 + a_2\lambda_2 y_2 + \cdots + a_t\lambda_t y_t = \boldsymbol{0} \\ a_1\lambda_1^2 y_1 + a_2\lambda_2^2 y_2 + \cdots + a_t\lambda_t^2 y_t = \boldsymbol{0} \\ \vdots \\ a_1\lambda_1^t y_1 + a_2\lambda_2^t y_2 + \cdots + a_t\lambda_t^t y_t = \boldsymbol{0} \end{cases}$

$\square$

(Cont.) 2nd Method.

- Let $\boldsymbol{x}_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{it} \end{pmatrix}$ for $i = 1, 2, \ldots, t$.
- Then $(a_1 x_{1i}, a_2 x_{2i}, \ldots, a_t x_{ti})^T$ satisfies that linear system, for all $i = 1, 2, \ldots, t$.
- While $\det \begin{pmatrix} \lambda_1 & \lambda_2 & \ldots & \lambda_t \\ \lambda_1^2 & \lambda_2^2 & \ldots & \lambda_t^2 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^t & \lambda_2^t & \ldots & \lambda_t^t \end{pmatrix} = \prod_{i=1}^t \lambda_i \prod_{1 \le i < j \le t} (\lambda_i - \lambda_j) \neq 0$. That is, that linear system has only trivial zero solution.
- Since $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_t$ are nonzero vectors, $a_1 = a_2 = \cdots = a_t = 0$. Contradiction.

$\square$

Exercise

Try to solve Exercise 6.2, 6.4, 6.5, 6.13.

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.

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?*

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 $n \geq 2$.

Page 3 Add 3 remarks for understanding some concepts and properties;

Page 6 Add 1 slide for "Conditions for diagonalizability";

Page 7 Add 1 additional question;

Page 11 Revise the solution for 2nd case;

Page 15 Revise the proof for Remark 6.2.5(2);

Page 16–18 Add 2 proofs for Remark 6.2.5(3);

Page 19 Add 3 additional questions;

Page 20 Add 2 additional questions.

Last modified: 15:28, March 23st, 2010.

Schedule of Today

- Any question about last tutorial
- Review concepts: Inner product, orthogonal, Gram-Schmidt process, projection,least squares solution, orthogonal matrix
- Tutorial: 5.9, 5.11, 5.15, 5.18, 5.24, 5.33(ab)
- Additional material: Orthogonal diagonalization

Concepts

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 $\cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)$.
- $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbb{R}^n$ are called orthogonal if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.
- A set S of vectors in $\mathbb{R}^n$ is called orthogonal if every pair of distinct vectors in S are orthogonal. (Orthogonal basis)
- If S is an orthogonal set of nonzero vectors in a vector space, then S is linearly independent. (By contrapositive)
- 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. (Orthonormal basis)
- Let V be a subspace of $\mathbb{R}^n$. 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.

Gram-Schmidt process

Let $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$ be a basis for a vector space V. Let

$$\begin{aligned}
v_1 &= u_1 \\
v_2 &= u_2 - \frac{u_2 \cdot v_1}{\|v_1\|^2} v_1 \\
v_3 &= u_3 - \frac{u_3 \cdot v_1}{\|v_1\|^2} v_1 - \frac{u_3 \cdot v_2}{\|v_2\|^2} v_2 \\
&\vdots \\
v_k &= u_k - \frac{u_k \cdot v_1}{\|v_1\|^2} v_1 - \frac{u_k \cdot v_2}{\|v_2\|^2} v_2 - \cdots - \frac{u_k \cdot v_{k-1}}{\|v_{k-1}\|^2} v_{k-1}
\end{aligned}$$

Then $\{v_1, v_2, \ldots, v_k\}$ is an orthogonal basis for V. Furthermore, let $w_i = \frac{v_i}{\|v_i\|}$ for $i = 1, 2, \ldots, k$. Then $\{w_1, w_2, \ldots, w_k\}$ is an orthogonal basis for V.

Projection

Let V be a subspace of $\mathbb{R}^n$, and $\boldsymbol{u}$ an arbitrary vector of $\mathbb{R}^n$.

- $\boldsymbol{u}$ can be written uniquely (by contrapositive) 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 projection of $\boldsymbol{u}$ onto V.
- If $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ is an orthogonal basis for V, then
$$\frac{\boldsymbol{u} \cdot \boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\boldsymbol{u}_1 + \frac{\boldsymbol{u} \cdot \boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2}\boldsymbol{u}_2 + \cdots + \frac{\boldsymbol{u} \cdot \boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}\boldsymbol{u}_k$$
is the projection of $\boldsymbol{u}$ 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.

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 $\|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|$.

- $\boldsymbol{x}$ is the least squares solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, iff $\boldsymbol{x}$ is the solution $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{p}$ where $\boldsymbol{p}$ is the projection of $\boldsymbol{b}$ onto the column space of $\boldsymbol{A}$, iff $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$.
- $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$ is always consistent:
$$\begin{aligned}\operatorname{rank}(\boldsymbol{A}^T\boldsymbol{A}|\boldsymbol{A}^T\boldsymbol{b}) &= \operatorname{rank}(\boldsymbol{A}^T(\boldsymbol{A}|\boldsymbol{b})) \leq \min\{\operatorname{rank}(\boldsymbol{A}^T), \operatorname{rank}(\boldsymbol{A}|\boldsymbol{b})\} \\ &= \min\{\operatorname{rank}(\boldsymbol{A}), \operatorname{rank}(\boldsymbol{A}|\boldsymbol{b})\} = \operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}^T\boldsymbol{A})\end{aligned}$$

Orthogonal 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 $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$.
- If λ is an eigenvalue of $\boldsymbol{A}$ which is an orthogonal matrix, then $|\lambda| = 1$: Since $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ and $\|\boldsymbol{A}\boldsymbol{x}\| = \|\boldsymbol{x}\|$, we have $|\lambda| = 1$.

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 (5.9)

Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ 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. $\square$

Exercise (5.11)

Let $\boldsymbol{u}_1 = (1,2,2,-1)$, $\boldsymbol{u}_2 = (1,1,-1,1)$, $\boldsymbol{u}_3 = (-1,1,-1,-1)$, $\boldsymbol{u}_4 = (-2,1,1,2)$.

(a) *Show that* $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\}$ *is an orthogonal set.*

(b) *Obtain an orthonormal set* S' *by normalizing* $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4$.

(c) *Is* S' *an orthonormal basis for* $\mathbb{R}^4$*?*

(d) *If* $\boldsymbol{w} = (0,1,2,3)$, *find* $(\boldsymbol{w})_S$ *and* $(\boldsymbol{w})_{S'}$.

(e) *Let* $V = \operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$. *Find all vectors that are orthogonal to* V.

(f) *Find the projection of* $\boldsymbol{w}$ *onto* V.

Proof and Solution.

(a) It is easy to check that $\boldsymbol{u}_i \cdot \boldsymbol{u}_j = 0$ for $i \neq j$.

(b) $S' = \left\{ \frac{1}{\sqrt{10}}(1,2,2,-1), \frac{1}{2}(1,1,-1,1), \frac{1}{2}(-1,1,-1,-1), \frac{1}{\sqrt{10}}(-2,1,1,2) \right\}$.

(c) Yes, since $|S'| = 4 = \dim(\mathbb{R}^4)$ and S' is orthonormal.

(d) $(\boldsymbol{w})_S = (\boldsymbol{w}\cdot\boldsymbol{u}_1, \boldsymbol{w}\cdot\boldsymbol{u}_2, \boldsymbol{w}\cdot\boldsymbol{u}_3, \boldsymbol{w}\cdot\boldsymbol{u}_4) = \left(\frac{3}{10}, \frac{1}{2}, -1, \frac{9}{10}\right)$ and
$(\boldsymbol{w})_{S'} = \left(\frac{3}{\sqrt{10}}, 1, -2, \frac{9}{\sqrt{10}} \right)$.

(e) A vector $\boldsymbol{v}$ is orthogonal to V if and only if $\boldsymbol{v} = t(-2,1,1,2)$ for some $t \in \mathbb{R}$, i.e. $\boldsymbol{v} \in \operatorname{span}\{(-2,1,1,2)\}$.

(f) $\boldsymbol{w} - \frac{\boldsymbol{w}\cdot\boldsymbol{u}_4}{\|\boldsymbol{u}_4\|}\frac{\boldsymbol{u}_4}{\|\boldsymbol{u}_4\|} = \left(\frac{9}{5}, \frac{1}{10}, \frac{11}{10}, \frac{6}{5}\right)$.

Exercise (5.15)

(a) *Find an orthonormal basis for the solution space of the equation* $x + y - z = 0$.

(b) *Find the projection of* $(1, 0, -1)$ *onto the plane* $x + y - z = 0$.

(c) *Extend the set obtained in Part (a) to an orthonormal basis for* $\mathbb{R}^3$.

Solution.

(a) A general solution to $x + y - z = 0$ is $\begin{cases} x = t - s \\ y = s \\ z = t \end{cases}$ where $t, s \in \mathbb{R}$. So $\{(-1, 1, 0), (1, 0, 1)\}$ is a basis for the solution space. Using Gram-Schmidt process, we transform this basis into an orthonormal basis $\left\{\frac{1}{\sqrt{2}}(-1, 1, 0), \frac{1}{\sqrt{6}}(1, 1, 2)\right\}$.

(b) $\frac{(1,0,-1)\cdot(-1,1,0)}{2}(-1, 1, 0) + \frac{(1,0,-1)\cdot(1,1,2)}{6}(1, 1, 2) = (\frac{1}{3}, -\frac{2}{3}, -\frac{1}{3})$.

(c) Since $(1, 1, -1)$ is orthogonal to the plane $x + y - z = 0$, it is orthogonal to the vectors in the basis obtained in (a). So $\left\{\frac{1}{\sqrt{2}}(-1, 1, 0), \frac{1}{\sqrt{6}}(1, 1, 2), \frac{1}{\sqrt{3}}(1, 1, -1)\right\}$ is an orthonormal basis for $\mathbb{R}^3$.

□

Remark

For a plane $ax + by + cz = d$ in $\mathbb{R}^3$, the vector (a, b, c) is called the normal vector of the plane, which is orthogonal to it. For more see Wikipedia.

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 $\frac{(5,3,1)\cdot(1,1,1)}{3}(1,1,1) = (3,3,3)$.
- When $p \neq 1$. It is easy to show $V = \text{span}\{(1,0,0),(0,1,1)\}$. 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 $\big((5,3,1)\cdot(1,0,0)\big)(1,0,0) + \frac{(5,3,1)\cdot(0,1,1)}{2}(0,1,1) = (5,2,2)$.

$\square$

Exercise (5.24)

Let $\boldsymbol{A} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, $\boldsymbol{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ *and* $\boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

(a) *Find the least squares solution to the linear system* $\boldsymbol{Ax} = \boldsymbol{b}$.

(b) *By the result in (a), compute the projection of* $\boldsymbol{b}$ *onto the column space of* $\boldsymbol{A}$.

Proof.

(a) By solving $\begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, we get the least squares solution is $x_1 = x_2 = x_3 = \frac{1}{3}$.

(b) $\boldsymbol{Ax} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ is the projection of $\boldsymbol{b}$ onto V, since $\boldsymbol{x}$ is the least squares solution to $\boldsymbol{Ax} = \boldsymbol{b}$.

$\square$

Exercise (5.33(ab))

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* $\mathrm{span}\{\boldsymbol{v}, \boldsymbol{w}\}$.

(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* $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}_k$.

(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* $\boldsymbol{A}\boldsymbol{B}$ *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 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.

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

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

$\square$

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) Choose $\boldsymbol{u} = e_i$, $\boldsymbol{v} = e_j$, 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 $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

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.

Let $\boldsymbol{v}$ be a solution of $\boldsymbol{Ax} = \boldsymbol{b}$, i.e. $\boldsymbol{Av} = \boldsymbol{b}$. Since $\boldsymbol{A}^T\boldsymbol{Av} = \boldsymbol{A}^T\boldsymbol{b}$, $\boldsymbol{v}$ is also a solution of $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$. Then

$$\begin{aligned}
\text{The solution set of } (\boldsymbol{Ax} = \boldsymbol{b}) &= \{\boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u} \in \text{nullspace of } (\boldsymbol{A})\} \\
&= \{\boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u} \in \text{nullspace of } (\boldsymbol{A}^T\boldsymbol{A})\} \\
&= \text{The solution set of } (\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b})
\end{aligned}$$

□

Exercise

Try to solve Exercise 5.30, 5.31, 5.32, 6.21, 6.27.

Exercise (Question 4 in Final of 2003-2004(II))

Let $\boldsymbol{W}$ *be a subspace of* $\mathbb{R}^n$ *and let* $\boldsymbol{W}^\perp = \{\boldsymbol{u} \in \mathbb{R}^n : \boldsymbol{u}$ *is orthogonal to* $\boldsymbol{W}\}$*. Then*

(i) $\boldsymbol{W}^\perp$ *is a subspace of* $\mathbb{R}^n$*;*

(ii) $\dim(\boldsymbol{W}) + \dim(\boldsymbol{W}^\perp) = n$*.*

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.*

Page 5 Add the proof for "Consistence of $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$";

Page 10 Add 1 remark for "Normal vector";

Page 13 Change "orthogonal" to "orthonormal" for part(cd);

Page 14 Revise the solution of part(h).

Last modified: 15:28, March 29th, 2010.

Schedule of Today

- Any question about last tutorial
- Review concepts
- Tutorial: 5.27, 5.33(cdef), 7.5, 7.11, 7.12, 7.14
- Additional material: 7.3, 7.13

Transition matrix

- Let $S = \{u_1, \dots, u_n\}$ and $T = \{v_1, \dots, v_n\}$ be two bases for the n dimensional vector space V. Then the transition matrix $\boldsymbol{P}$ from S to T is uniquely determined by the equation

$$(u_1, \dots, u_n) = (v_1, \dots, v_n)\boldsymbol{P} = (v_1, \dots, v_n)([u_1]_T, \dots, [u_n]_T).$$

- For any vector $\boldsymbol{w}$, $[\boldsymbol{w}]_S$ and $[\boldsymbol{w}]_T$ are the coordinates in S and T coordinate systems, respectively. Then

$$(v_1, \dots, v_n)[w]_T = w = (u_1, \dots, u_n)[w]_S = (v_1, \dots, v_n)P[w]_S,$$

hence

$$[w]_T = \boldsymbol{P}[w]_S.$$

Linear transformation

- A linear transformation is a mapping $T: \mathbb{R}^n \to \mathbb{R}^m$ of the form

$$T\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{for all} \quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

  where a_{ij} is a real number for $1 \leq i \leq m$, $1 \leq j \leq n$. The matrix $(a_{ij})_{m\times n}$ is called the standard matrix for T.

- How to find the standard matrix for T: solving the equation

$$T(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n) = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m)\boldsymbol{A}.$$

- A mapping $T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if and only if

$$T(a\boldsymbol{u} + b\boldsymbol{v}) = aT(\boldsymbol{u}) + bT(\boldsymbol{v}) \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n, a, b \in \mathbb{R}.$$

  This result can be used in the Final Exam.

- If $n = m$, T is also called a linear operator on $\mathbb{R}^n$.

Properties

- 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 again by $\boldsymbol{BA}$.

Rank and Kernal

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and $\boldsymbol{A}$ the standard matrix for T.

Def: The range of T is the set of images of T, i.e.

$$\mathrm{R}(T) = \{T(\boldsymbol{u}) | \boldsymbol{u} \in \mathbb{R}^n\} \subset \mathbb{R}^m.$$

Thm: $\mathrm{R}(T) =$ the column space of $\boldsymbol{A}$.

Def: The dimension of $\mathrm{R}(T)$ is called the rank of T and is denoted by $\mathrm{rank}(T)$.

Thm: $\mathrm{rank}(T) = \mathrm{rank}(\boldsymbol{A})$.

Def: The kernal of T is the set of vectors in $\mathbb{R}^n$ whose image is the zero vector in $\mathbb{R}^m$, i.e.

$$\mathrm{Ker}(T) = \{\boldsymbol{u} | T(\boldsymbol{u}) = \boldsymbol{0}\} \subset \mathbb{R}^n.$$

Thm: $\mathrm{Ker}(T) =$ the nullspace of $\boldsymbol{A}$.

Def: The dimension of $\mathrm{Ker}(T)$ is called the nullity of T and is denoted by $\mathrm{nullity}(T)$.

Thm: $\mathrm{rank}(T) + \mathrm{nullity}(T) = \mathrm{rank}(\boldsymbol{A}) + \mathrm{nullity}(\boldsymbol{A}) = n$.

- For 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 (5.27)

Suppose an $x'y'$-coordinate system is obtained from the xy-coordinate system by an anti-clockwise rotation through an angle $\theta = \frac{\pi}{3}$.

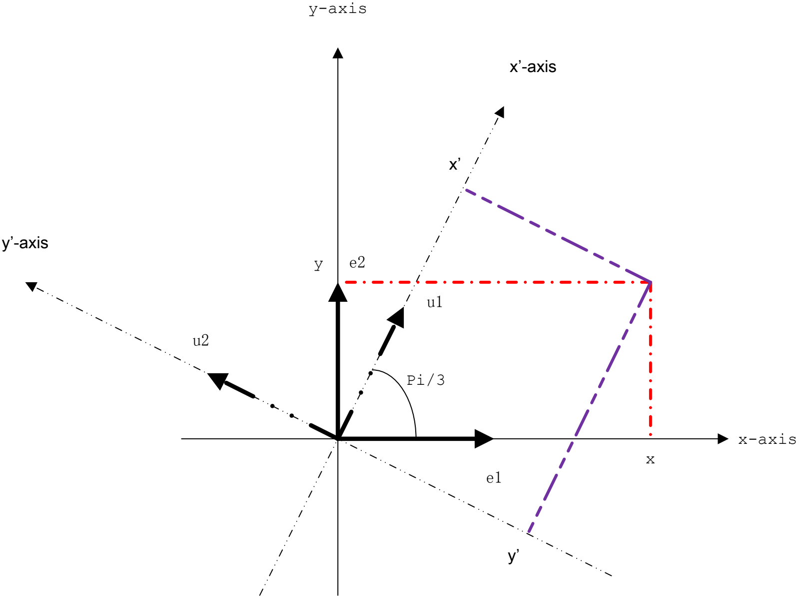
(a) *Let P be the point such that its xy-coordinates are $(2, 1)$. Find the $x'y'$-coordinates of P.*

(b) *Let Q be the point such that its $x'y'$-coordinates are $(2, 1)$. Find the xy-coordinates of Q.*

(c) *Let l be the line $x + y = 1$. Write down the equation of l using the $x'y'$-coordinates.*

Solution.

Let $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ be the unit vectors such that $\boldsymbol{e}_1$ is in the direction of the x-axis and $\boldsymbol{e}_2$ is in the direction of the y-axis. Also let $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ be the unit vectors such that $\boldsymbol{u}_1$ is in the direction of the x'-axis and $\boldsymbol{u}_2$ is in the direction of the y'-axis. □


y-axis
x'-axis
x'
y'-axis
y
e2
u1
u2
Pi/3
x-axis
x
e1
y'

Solution.

The transition matrix from $\{u_1, u_2\}$ to $\{e_1, e_2\}$ is:

$$\boldsymbol{P} = ([\boldsymbol{u}_1]_{\{e_1,e_2\}}, [\boldsymbol{u}_2]_{\{e_1,e_2\}}) = \begin{pmatrix} \cos(\pi/3) & -\sin(\pi/3) \\ \sin(\pi/3) & \cos(\pi/3) \end{pmatrix} = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}.$$

Since $(u_1, u_2) = (e_1, e_2)\boldsymbol{P}$ and $[w]_{\{e_1,e_2\}} = \boldsymbol{P}[w]_{\{u_1,u_2\}}$, we have:

(a) $\begin{pmatrix} x' \\ y' \end{pmatrix} = [\boldsymbol{w}]_{\{u_1,u_2\}} = \boldsymbol{P}^{-1}[\boldsymbol{w}]_{\{e_1,e_2\}} = \boldsymbol{P}^{-1}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1+\sqrt{3}/2 \\ 1/2-\sqrt{3} \end{pmatrix}.$

(b) $\begin{pmatrix} x \\ y \end{pmatrix} = [\boldsymbol{w}]_{\{e_1,e_2\}} = \boldsymbol{P}[\boldsymbol{w}]_{\{u_1,u_2\}} = \boldsymbol{P}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1-\sqrt{3}/2 \\ 1/2+\sqrt{3} \end{pmatrix}.$

(c) Since $\begin{pmatrix} x \\ y \end{pmatrix} = \boldsymbol{P}\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x'/2-\sqrt{3}y'/2 \\ \sqrt{3}x'/2+y'/2 \end{pmatrix}$, and $x+y=1$, we have $x'/2-\sqrt{3}y'/2+\sqrt{3}x'/2+y'/2=1$, i.e.

$$(1+\sqrt{3})x' + (1-\sqrt{3})y' = 2.$$

$\square$

Remark

There are some differences on the concept "transition matrix" between the textbook and some reference books, you have better to remember the definition of textbook.

Exercise (5.33(cdef))

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

(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* $\boldsymbol{A}\boldsymbol{B}$ *is an orthogonal matrix.*

(g) *If the columns of a square matrix* $\boldsymbol{A}$ *form an orthonormal set, then the rows of* $\boldsymbol{A}$ *also form an orthonormal set.*

(g)' *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.

(c) True. By definition.

(c') False. Choose $\boldsymbol{c} = \mathbf{0}$

(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.

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

$\square$

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.*

$$(S \circ T)(\boldsymbol{u}) = \boldsymbol{u} \quad \textit{for all } \boldsymbol{u} \in \mathbb{R}^n,$$

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* $T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x \\ y \end{pmatrix}$ *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*.*

Proof.

(a) (i) Since $T(\boldsymbol{u}) = \boldsymbol{u}$ for all $\boldsymbol{u} \in \mathbb{R}$, $T(T(\boldsymbol{u})) = T(\boldsymbol{u}) = \boldsymbol{u}$. 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 (7.11)

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)
- Let $\boldsymbol{u} \in \mathrm{Ker}(S)$, i.e. $S(\boldsymbol{u}) = \mathbf{0}$.
- Then $T \circ S(\boldsymbol{u}) = T(S(\boldsymbol{u})) = T(\mathbf{0}) = \mathbf{0}$ and hence $\boldsymbol{u} \in \mathrm{Ker}(T \circ S)$.
- Thus $\mathrm{Ker}(S) \subset \mathrm{Ker}(T \circ S)$.

(b)
- Let $\boldsymbol{v} \in \mathrm{R}(T \circ S)$, i.e. there exists $\boldsymbol{u} \in \mathbb{R}^n$ such that $\boldsymbol{v} = T \circ S(\boldsymbol{u})$.
- Put $\boldsymbol{w} = S(\boldsymbol{u}) \in \mathbb{R}^m$. Then $\boldsymbol{v} = T(S(\boldsymbol{u})) = T(\boldsymbol{w})$.
- This means that $\boldsymbol{v} \in \mathrm{R}(T)$. Thus $\mathrm{R}(T \circ S) \subset \mathrm{R}(T)$.

$\square$

Exercise (7.12)

Let $\boldsymbol{n}$ be a unit vector in $\mathbb{R}^n$. Define $P : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$P(\boldsymbol{x}) = \boldsymbol{x} - (\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^n.$$

(a) *Show that P is a linear transformation and find the standard matrix for P.*

(b) *Prove that $P \circ P = P$.*

Proof.

(a) For any $\boldsymbol{x} \in \mathbb{R}^n$, $P(\boldsymbol{x}) = \boldsymbol{x} - (\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} = \boldsymbol{I}_n\boldsymbol{x} - \boldsymbol{n}\boldsymbol{n}^T\boldsymbol{x} = (\boldsymbol{I}_n - \boldsymbol{n}\boldsymbol{n}^T)\boldsymbol{x}$. So P is a linear transformation and the standard matrix for P is $\boldsymbol{I} - \boldsymbol{n}\boldsymbol{n}^T$.

(b) Since for all $\boldsymbol{x} \in \mathbb{R}^n$,

$$\begin{aligned}(P \circ P)(\boldsymbol{x}) &= P(\boldsymbol{x} - (\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n}) \\ &= \boldsymbol{x} - (\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} - \{\boldsymbol{n} \cdot [\boldsymbol{x} - (\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n}]\}\boldsymbol{n} \\ &= \boldsymbol{x} - (\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} - \{(\boldsymbol{n} \cdot \boldsymbol{x}) - (\boldsymbol{n} \cdot \boldsymbol{x})(\boldsymbol{n} \cdot \boldsymbol{n})\}\boldsymbol{n} \\ &= \boldsymbol{x} - (\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} \\ &= P(\boldsymbol{x})\end{aligned}$$

$P \circ P = P$.

□

Remark

The linear operator P is called projection, whose standard matrix is idempotent.

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 $T(\boldsymbol{v}) = \mathbf{0}$.*

(c) *Find all linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $T \circ T = T$.*

Proof.

(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

$$T(\boldsymbol{u}) = T(T(\boldsymbol{x})) = (T \circ T)(\boldsymbol{x}) = T(\boldsymbol{x}) = \boldsymbol{u}.$$

(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

$$T(\boldsymbol{v}) = T(T(\boldsymbol{y}) - \boldsymbol{y}) = (T \circ T)(\boldsymbol{y}) - T(\boldsymbol{y}) = T(\boldsymbol{y}) - T(\boldsymbol{y}) = \mathbf{0}.$$

□

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 $\boldsymbol{x}$ a 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: $\lambda_1 = \lambda_2 = 0$. By (a), we cannot find a nonzero vector $\boldsymbol{u}$, such that $T(\boldsymbol{u}) = \boldsymbol{u}$; Otherwise T has a eigenvalue 1. Then T is the zero transformation.
- Case 2: $\lambda_1 = \lambda_2 = 1$. By (b), we cannot find a nonzero vector $\boldsymbol{v}$, such that $T(\boldsymbol{v}) = \boldsymbol{0}$; Otherwise T has a eigenvalue 0. 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 $\boldsymbol{P} = \begin{pmatrix} 1 & b \\ c & d \end{pmatrix}$, 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
$$\boldsymbol{A} = \boldsymbol{0}_2, \boldsymbol{I}_2, \begin{pmatrix} r & s \\ t & 1-r \end{pmatrix} \quad \text{where } st = r(1-r).$$

$\square$

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.

"$\Leftarrow$": Suppose

$$T(a\boldsymbol{u} + b\boldsymbol{v}) = aT(\boldsymbol{u}) + bT(\boldsymbol{v}) \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n, a, b \in \mathbb{R}.$$

Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ be the canonical basis for $\mathbb{R}^n$ and let $\boldsymbol{A}$ be the $m \times n$ matrix $\begin{pmatrix} T(\boldsymbol{e}_1) & T(\boldsymbol{e}_2) & \cdots & T(\boldsymbol{e}_n) \end{pmatrix}$. For any $\boldsymbol{u} = (u_1, u_2, \dots, u_n)^T \in \mathbb{R}^n$, $\boldsymbol{u} = u_1\boldsymbol{e}_1 + u_2\boldsymbol{e}_2 + \cdots + u_n\boldsymbol{e}_n$. Then we have

$$\begin{aligned} T(\boldsymbol{u}) &= u_1 T(\boldsymbol{e}_1) + u_2 T(\boldsymbol{e}_2) + \cdots + u_n T(\boldsymbol{e}_n) \\ &= \begin{pmatrix} T(\boldsymbol{e}_1) & T(\boldsymbol{e}_2) & \cdots & T(\boldsymbol{e}_n) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \boldsymbol{A}\boldsymbol{u} \end{aligned}$$

Thus T is a linear transformation.

Exercise (7.13)

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.*

(b) *Show that $F \circ F$ is the identity transformation.*

(c) *Show that the standard matrix for F is an orthogonal matrix.*

Proof.

(a) Similar to Exercise 7.12, for any $\boldsymbol{x} \in \mathbb{R}^n$, $F(\boldsymbol{x}) = \boldsymbol{I}_n\boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} = (\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T)\boldsymbol{x}$. So F is a linear transformation and the standard matrix for F is $\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T$.

□

Proof of part (bc).

(b) Since for all $\boldsymbol{x} \in \mathbb{R}^n$,

$$\begin{aligned}(F \circ F)(\boldsymbol{x}) = F(F(\boldsymbol{x})) &= F(\boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n}) \\ &= \boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} - 2\{\boldsymbol{n} \cdot [\boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n}]\}\boldsymbol{n} \\ &= \boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} - 2\{(\boldsymbol{n} \cdot \boldsymbol{x}) - 2(\boldsymbol{n} \cdot \boldsymbol{x})(\boldsymbol{n} \cdot \boldsymbol{n})\}\boldsymbol{n} \\ &= \boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} - 2\{-(\boldsymbol{n} \cdot \boldsymbol{x})\} \cdot \boldsymbol{n} = \boldsymbol{x},\end{aligned}$$

$F \circ F$ is the identity transformation.

(b') Alternatively, $(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T)^2 = (\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T)(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T) = \boldsymbol{I} - 4\boldsymbol{n}\boldsymbol{n}^T + 4\boldsymbol{n}\boldsymbol{n}^T\boldsymbol{n}\boldsymbol{n}^T = \boldsymbol{I}$ since $\boldsymbol{n}$ is a unit vector.

(c) Note that $(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T)^T = \boldsymbol{I} - 2(\boldsymbol{n}\boldsymbol{n}^T)^T = \boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T$. Thus

$$(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T)(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T)^T = (\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^T)^2 = \boldsymbol{I}$$

by (b). The standard matrix is an orthogonal matrix.

$\square$

Remark

F is a reflection about the hyperplane which is orthogonal to $\boldsymbol{n}$.

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 $\{\boldsymbol{0}\}$.

Exercise (Question 3 in Final 2004-2005(II))

Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ 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 $T(\boldsymbol{e}_n) = \boldsymbol{0}$. Find all the eigenvalues and eigenvectors of $\boldsymbol{A}$, where $\boldsymbol{A}$ is the standard matrix for T.

Solution.

$$T(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) = (\boldsymbol{e}_2, \ldots, \boldsymbol{e}_n, \boldsymbol{0}) = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)\begin{pmatrix} 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 1 & 0 \end{pmatrix}. \qquad \square$$

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*

$$\mathrm{Ker}(T) \cap \mathrm{R}(T) = \{\mathbf{0}\}.$$

Page 4 Add 1 remark that "Exercise 7.3 can be used in the Final Exam";

Page 6 Add 1 remark that "For 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";

Page 9 Add 1 remark that "There are some differences on the concept "transition matrix" between the textbook and some reference books, you have better to remember the definition of textbook";

Page 17 Revise the solution of Exercise 7.14(c);

Page 18 Revise 1 typo.

Last modified: 13:35, April 06th, 2010.

Information of Final Exam

- Time: April 24th, 09:00–11:00;
- Venue: MPSH2;
- Close book with 2 helpsheets (double-sided, A4 size);
- Consultation: till April 24th, 09:00
  - Office: S17-06-14.
  - Mobile: 9053-5550.
  - Email: xiangsun@nus.edu.sg.
- Good luck!

Schedule of Today

- Any question about last tutorial
- Review concepts
- Tutorial:
- Additional material

Matrix

- Definition: A matrix is a rectangular array of numbers. A number can be regarded as a 1×1 matrix, and a 1×1 matrix also be can regarded as a number as you want;
- Notation: $\boldsymbol{A} = (a_{ij})_{m \times n}$; for simpleness, we use $\mathbb{R}^{m \times n}$ to denote the sets of all $m \times n$ matrices.
- Matrix operations:
  - addition: $\boldsymbol{A} + \boldsymbol{B} = (a_{ij} + b_{ij})_{m \times n}$;
  - scalar multiplication: $\lambda \boldsymbol{A} = (\lambda a_{ij})_{m \times n}$;
  - matrix multiplication: $\boldsymbol{A}_{m \times n} \boldsymbol{B}_{n \times p} = \boldsymbol{C}_{m \times p}$, where $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$;
  - transpose: $\boldsymbol{A}^T = (a_{ji})_{n \times m}$;
  - trace: $\operatorname{tr}(\boldsymbol{A}) = a_{11} + a_{22} + \cdots + a_{nn}$.
- Block matrix:

$$\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1q} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{A}_{p1} & \boldsymbol{A}_{p2} & \cdots & \boldsymbol{A}_{pq} \end{pmatrix}$$

  the operations of block matrices are similar to the matrix operations. See block matrix@Wiki.

Determinant

- Definition: Let $\boldsymbol{A} = (a_{ij}) \in \mathbb{R}^{n\times n}$. Let $\boldsymbol{M}_{ij}$ be an matrix obtained from $\boldsymbol{A}$ by deleting the i-th row and the j-th 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 $A_{ij} = (-1)^{i+j}\det(M_{ij})$, which is called the (i,j)-cofactor of $\boldsymbol{A}$;
- General properties;
- $\det(\boldsymbol{A}^T) = \det(\boldsymbol{A})$; $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$;
- $\det\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1p} \\ \boldsymbol{0} & \boldsymbol{A}_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{A}_{p-1,p} \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{A}_{pp} \end{pmatrix} = \det(\boldsymbol{A}_{11})\det(\boldsymbol{A}_{22})\cdots\det(\boldsymbol{A}_{pp})$;
- Cofactor expansion: $\det(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ij}A_{ij} = \sum_{j=1}^{n} a_{ij}A_{ij}$;
- Computation:
  - Using elementary operations;
  - Using cofactor expansion;
  - Decompose the matrix as a product of some simple matrices.

Adjoint matrix

- Let $\boldsymbol{A} = (a_{ij}) \in \mathbb{R}^{n\times n}$. Then the 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 $\boldsymbol{A}$.
- $\boldsymbol{A}\operatorname{adj}(\boldsymbol{A}) = \operatorname{adj}(\boldsymbol{A})\boldsymbol{A} = \det(\boldsymbol{A})\boldsymbol{I}_n$, this equation holds for any invertible matrix $\boldsymbol{A}$ and any singular matrix $\boldsymbol{A}$;
- $\operatorname{adj}(\boldsymbol{AB}) = \operatorname{adj}(\boldsymbol{B})\operatorname{adj}(\boldsymbol{A})$.

Inverse

- Definition: Let $\boldsymbol{A} = (a_{ij}) \in \mathbb{R}^{n\times n}$ matrix. Then $\boldsymbol{A}$ is said to be invertible if there exists a square matrix $\boldsymbol{B}$ of order n such that $\boldsymbol{AB} = \boldsymbol{I}$ and $\boldsymbol{BA} = \boldsymbol{I}$. 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;
- A matrix $\boldsymbol{A}$ is invertible iff $\det(\boldsymbol{A}) \neq 0$; moreover, if $\det(\boldsymbol{A}) \neq 0$, then $\boldsymbol{A}^{-1} = \frac{1}{\det(\boldsymbol{A})}\operatorname{adj}(\boldsymbol{A})$.
- $(\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T$, $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$;
- For block matrix:

$$\begin{pmatrix} \boldsymbol{A}_{11} & * & \cdots & * \\ \boldsymbol{0} & \boldsymbol{A}_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{A}_{pp} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{A}_{11}^{-1} & * & \cdots & * \\ \boldsymbol{0} & \boldsymbol{A}_{22}^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{A}_{pp}^{-1} \end{pmatrix};$$

- Computation:
  - Elementary row operations: $(\boldsymbol{A}|\boldsymbol{I}) \to (\boldsymbol{I}|\boldsymbol{A}^{-1})$;
  - Adjoint: $\boldsymbol{A}^{-1} = \frac{1}{\det(\boldsymbol{A})}\operatorname{adj}(\boldsymbol{A})$;
  - Solving the linear system: $\boldsymbol{AX} = \boldsymbol{I}$;
  - Decompose the matrix as a product of some simple matrices.

Rank

- 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

$$\begin{aligned}\operatorname{rank}(\boldsymbol{A}) &= \#\text{ non-zero rows of } \boldsymbol{R} = \#\text{ leading entries of } \boldsymbol{R} = \#\text{ pivot columns of } \boldsymbol{R}\\ &= \text{largest } \#\text{ of L.I. rows in } \boldsymbol{A} = \text{largest } \#\text{ of L.I. columns in } \boldsymbol{A}\\ &= \text{largest size of invertible submatrices of } \boldsymbol{A}\end{aligned}$$

- $\boldsymbol{A} \in \mathbb{R}^{m\times n}$. The solution space of the homogeneous system of linear equations $\boldsymbol{A}\boldsymbol{x} = 0$ is called nullspace of $\boldsymbol{A}$, and $\dim(\text{nullspace of } \boldsymbol{A})$ is called the nullity of $\boldsymbol{A}$, denoted by $\operatorname{nullity}(\boldsymbol{A})$.
- $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, then $\operatorname{rank}(\boldsymbol{A}) + \operatorname{nullity}(\boldsymbol{A}) = (\#\text{ columns of } \boldsymbol{A}) = n$.
- $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, then $\operatorname{rank}(\boldsymbol{A}) \leq \min\{m, n\}$.
- $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, then $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}^T)$.
- $\boldsymbol{B}$ is a submatrix of $\boldsymbol{A}$, then $\operatorname{rank}(\boldsymbol{B}) \leq \operatorname{rank}(\boldsymbol{A})$.

(1) $\mathrm{rank}(\boldsymbol{AB}) \leq \min\{\mathrm{rank}(\boldsymbol{A}), \mathrm{rank}(\boldsymbol{B})\}$. 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 $\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{PA}) = \mathrm{rank}(\boldsymbol{AQ}) = \mathrm{rank}(\boldsymbol{PAQ})$. By (1) or see Exercise 4.22.

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

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

(2c) $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\mathrm{rank}(\boldsymbol{A}) = r$, 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 $\mathrm{rank}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix} \geq \mathrm{rank}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix}$. Def.

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

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

(4a) $\max\{\mathrm{rank}(\boldsymbol{A}), \mathrm{rank}(\boldsymbol{B})\} \leq \mathrm{rank}(\boldsymbol{A}|\boldsymbol{B}) \leq \mathrm{rank}(\boldsymbol{A}) + \mathrm{rank}(\boldsymbol{B})$, $\mathrm{rank}\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix} \leq \mathrm{rank}(\boldsymbol{A}) + \mathrm{rank}(\boldsymbol{B})$. By (4).

(4c) $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, then $\mathrm{rank}(\boldsymbol{A}) + \mathrm{rank}(\boldsymbol{I}_n + \boldsymbol{A}) \geq n$. By (4).

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

Vector space

- Vector space: 8 axioms; Subspace: closed under addition and scalar multiplication;
- Linear independent: Let $S = \{u_1, u_2, \ldots, u_k\} \subset \mathbb{R}^n$. S is called a linearly independent set and $u_1, u_2, \ldots, u_k$ are said to be linearly independent if the equation
$$c_1 u_1 + c_2 u_2 + \cdots + c_k u_k = 0$$
has only trivial solution, where $c_1, c_2, \ldots, c_k$ are variables. Otherwise, S is called a linearly dependent set and $u_1, u_2, \ldots, u_k$ are said to be linearly dependent, i.e. there exist real numbers $a_1, a_2, \ldots, a_k$, not all of them are zero, such that $a_1 u_1 + a_2 u_2 + \cdots + a_k u_k = 0$.
- Let $S = \{u_1, u_2, \ldots, u_k\}$ be a subset of a vector space V. Then S is called a basis for V if
  - S is linearly independent;
  - S spans V.
- 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.

Transition matrix

- Let $S = \{u_1, \dots, u_n\}$ and $T = \{v_1, \dots, v_n\}$ be two bases for the n dimensional vector space V. Then the transition matrix $\boldsymbol{P}$ from S to T is uniquely determined by the equation

$$(u_1, \dots, u_n) = (v_1, \dots, v_n)\boldsymbol{P} = (v_1, \dots, v_n)([u_1]_T, \dots, [u_n]_T).$$

- For any vector $\boldsymbol{w}$, $[\boldsymbol{w}]_S$ and $[\boldsymbol{w}]_T$ are the coordinates in S and T coordinate systems, respectively. Then

$$(v_1, \dots, v_n)[w]_T = w = (u_1, \dots, u_n)[w]_S = (v_1, \dots, v_n)P[w]_S,$$

hence

$$[w]_T = \boldsymbol{P}[w]_S.$$

Linear transformation

- A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if and only if
$$T(a\boldsymbol{u} + b\boldsymbol{v}) = aT(\boldsymbol{u}) + bT(\boldsymbol{v}) \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n, a, b \in \mathbb{R}.$$
- Standard matrix:
$$\boldsymbol{A} = T(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n).$$
- If $n = m$, T is also called a linear operator on $\mathbb{R}^n$.
- Operations:
  - addition: $(T + S)(\boldsymbol{u}) = T(\boldsymbol{u}) + S(\boldsymbol{u})$;
  - scalar multiplication: $(\lambda T)(\boldsymbol{u} = \lambda \cdot T(\boldsymbol{u})$;
  - composition: $(S \circ T)(\boldsymbol{u}) = S(T(\boldsymbol{u}))$.

Def: The range of T is the set of images of T, i.e. $\mathrm{R}(T) = \{T(\boldsymbol{u}) | \boldsymbol{u} \in \mathbb{R}^n\} \subset \mathbb{R}^m$.

Thm: $\mathrm{R}(T)$ = the column space of $\boldsymbol{A}$.

Def: The dimension of $\mathrm{R}(T)$ is called the rank of T and is denoted by $\mathrm{rank}(T)$.

Thm: $\mathrm{rank}(T) = \mathrm{rank}(\boldsymbol{A})$.

Def: The kernal of T is the set of vectors in $\mathbb{R}^n$ whose image is the zero vector in $\mathbb{R}^m$, i.e. $\mathrm{Ker}(T) = \{\boldsymbol{u} | T(\boldsymbol{u}) = \boldsymbol{0}\} \subset \mathbb{R}^n$.

Thm: $\mathrm{Ker}(T)$ = the nullspace of $\boldsymbol{A}$.

Def: The dimension of $\mathrm{Ker}(T)$ is called the nullity of T and is denoted by $\mathrm{nullity}(T)$.

Thm: $\mathrm{rank}(T) + \mathrm{nullity}(T) = \mathrm{rank}(\boldsymbol{A}) + \mathrm{nullity}(\boldsymbol{A}) = n$.

Inner product space

- Inner product;
- Orthogonal, orthonormal;
- Basis $\xrightarrow{\text{Gram-Schmidt}}$ orthogonal basis $\xrightarrow{\text{normalizing}}$ orthonormal basis;
- Projection.

Linear system

- $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent iff $\mathrm{rank}(\boldsymbol{A}|\boldsymbol{b}) = \mathrm{rank}(\boldsymbol{A})$.
- The general solution of $\boldsymbol{Ax} = \boldsymbol{b}$ is the general solution of $\boldsymbol{Ax} = \boldsymbol{0}$ + a special solution of $\boldsymbol{Ax} = \boldsymbol{b}$.
- The number of parameters of $\boldsymbol{Ax} = \boldsymbol{b}$ is $n - \mathrm{rank}(\boldsymbol{A})$.
- Least squares solution: $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$.

Eigenvalue

- 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 some invertible matrix, then $\boldsymbol{A}$ and $\boldsymbol{B}$ have same eigenvalues. While the converse is not necessarily true. (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 $\boldsymbol{A}$ has n eigenvalues $\{\lambda_i\}_{i=1}^n$, then $\mathrm{tr}(\boldsymbol{A}) = \sum_{i=1}^n \lambda_i$, $\det(\boldsymbol{A}) = \prod_{i=1}^n \lambda_i$. (Exercise 6.2(a))

Eigenvalue

- 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 $\boldsymbol{A}$,

$$\text{the algebraic multiplicity of } \lambda \geq \text{ the geometric multiplicity of } \lambda \geq 1.$$

Conditions for diagonalizability

- The square matrix $\boldsymbol{A}$ of order n is diagonalizable iff $\boldsymbol{A}$ has n linearly independent eigenvectors.
- If the square matrix $\boldsymbol{A}$ of order n 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.

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 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* $\mathbb{R}^3$.

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.

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 $\boldsymbol{B}$.

Thank you