

# Algebra

Instructor: Huong Tran

Email address: huong.ttt@vgu.edu.vn

Office: A109, Binh Duong campus

Office hours: 9.00-11.00 Monday mornings

Slides are partly prepared by Prof. Dr. Christina Anderson

January 7, 2019

# Lecture 1: Introduction and vector operation review

# Contents I

## 1. Linear algebra: $\sim$ 10-11 lectures

1.1 Review: Vector operations in 2 and 3 dimensional spaces

1.2 System of linear equations (LES)

- ▶ Solution existence of a LES
- ▶ LES and associated augmented matrices
- ▶ Gauss and Gauss-Jordan elimination method

1.3 Matrix Algebra

- ▶ Terminologies
- ▶ Operations with matrices
- ▶ The inverse of a matrix
- ▶ Partitioned matrix and matrix factorization (self-study)

1.4 Determinants

1.5 Vector spaces

1.6 Linear transformations

1.7 Coordinate systems and basic transition matrix

# Contents II

## 2. General algebra: $\sim$ 9-10 lectures

2.1 Propositional logic: Lehman et.al.'s (Chapter 3), or Rosen's (Chapter 1).

2.2 Set theory review: Rosen's textbook

2.3 Relations: Rosen's textbook (chapter 9)

- ▶ Binary relations
- ▶ Popular properties on relations: reflexivity, symmetry, transitivity, anti-symmetry.
- ▶ Closures of a relation

2.4 Group theory

- ▶ Basic definitions
- ▶ Generators for a group and cyclic groups
- ▶ Cosets and Lagrange's theorem
- ▶ Homomorphism and isomorphism
- ▶ Quotient groups

# References

- ▶ Linear algebra
  - 1. D. Lay, Linear Algebra and Its Applications, Pearson New International Edition, Pearson, 2014 (primary).
  - 2. Gilbert Strang, Linear Algebra and Its Applications, Fourth edition, Brooks/Cole Cengage Learning, 2006.
  - 3. Serge Lang, Introduction to Linear Algebra, Second edition, Springer.
- ▶ General algebra
  - 1. Kenneth Rosen, Discrete Mathematics and its applications, Mc Graw Hill education, 2013 ([Logic, relations, group](#)).
  - 2. Eric Lehman, F.T. Leighton, and A. R. Meyer, Mathematics for Computer Science, 2017 ([Logic](#)).
  - 3. Joseph A. Gallian, Contemporary Abstract Algebra, Cengage learning, 2017 ([Groups](#))

## Learning outcomes

1. Acquired with essential concepts, structures and methods of propositional algebra, general algebra and linear algebra. Particularly, well-acquired with basic algebraic structures necessary for the comprehension of formal structures in CS;
2. Have the ability to independently develop abstract concepts and to acquire basis techniques of algebra;
3. Acquired for analytical thinking, development of methodological expertise, handling abstract methods, structures and models.

## Attendance and exam

1. To enable to pass the final exam, you are recommended to
  - ▶ attend the class at least 70% number of contact hours
  - ▶ do exercises/homework as much as possible
2. To get further point of views as well as to understand applications of the subject in the computer science field, please spend your time for reading textbooks.
3. Exam duration: 90 minutes with 7-12 questions. You are allowed to bring a two-sided A4 written with any contents. Pocket calculator is Not allowed.

# Linear Algebra

Lecture 1: Vector operations in 2,3-dimensional spaces

# Outline

- ▶ Geometric representation of vectors
- ▶ Coordinate representation of vectors
- ▶ Operations on vectors
- ▶ Inner/dotted product
- ▶ Length/norm of vectors
- ▶ Orthogonal vectors
- ▶ Angle between two vectors

# Graphic representation of vectors

## Definition (Vector)

$A$ ,  $B$  are points in the plane (or in the space). The vector  $\vec{AB}$  is the directed segment from  $A$  to  $B$ .

## Remark

Some variables, such as length, temperature, time are totally determined by just one number. These variables are so-called scalars. Other variables, such as speed, force need both a magnitude and a direction to be determined completely. These variables are vectors.

**Examples:** On the whiteboard

## Definition (Equality of vectors)

Two vectors are equal if they have the same direction and the same length.

## Notations

- ▶ Vectors are parallel:  $\vec{AB} \uparrow\uparrow \vec{CD}$
- ▶ Vectors are antiparallel:  $\vec{AB} \uparrow\downarrow \vec{CD}$
- ▶ Vectors:  $\vec{a}, \vec{b}, \dots$
- ▶ Scalars:  $\alpha, \beta, \gamma, \dots, \lambda, \mu, \dots$

## Definition (Sum of the two vectors $\vec{a}$ and $\vec{b}$ )

If we shift  $\vec{b}$  in a parallel manner, such that its origin corresponds to the end point of  $\vec{a}$ , then  $\vec{a} + \vec{b}$  starts at the origin of  $\vec{a}$  and ends at the end of  $\vec{b}$ .

$$\vec{a} + \vec{b} = \vec{AB} + \vec{BC} = \vec{AC}$$

## Theorem

1. *Commutative law:*

$$\vec{a} + \vec{b} = \vec{b} + \vec{a}$$

2. *Associative law:*

$$(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$$

Proof of 1: On the whiteboard

Proof of 2: On the whiteboard

$$\vec{a} + \vec{b} = \vec{AC}$$

$$(\vec{a} + \vec{b}) + \vec{c} = \vec{AC} + \vec{CD} = \vec{AD}$$

$$\vec{b} + \vec{c} = \vec{BD}$$

$$\vec{a} + (\vec{b} + \vec{c}) = \vec{AB} + \vec{BD} = \vec{AD} = (\vec{a} + \vec{b}) + \vec{c}$$

Thus:

$$(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$$

## Definition (Special vectors)

1. Zero vector:

$$\vec{0} = \vec{AA}$$

2. Opposite vector to  $\vec{a} = \vec{AB}$ :

$$-\vec{a} = \vec{BA}$$

## Remark

$$\vec{a} + (-\vec{a}) = \vec{AB} + \vec{BA} = \vec{0}$$

## Definition

Let  $\vec{a}, \vec{b}$  be vectors. Then

$$\vec{a} - \vec{b} = \vec{a} + (-\vec{b})$$

## Definition (Multiplication with a scalar)

Given a vector  $\vec{a}$  and a scalar  $\lambda$ . Then, the multiplication  $\vec{a}$  with a scalar  $\lambda$  is a vector with the following properties:

1. Length  $(\lambda \cdot \vec{a}) = |\lambda| \cdot \text{length } (\vec{a})$
2. For  $\lambda > 0$  is  $\lambda \cdot \vec{a} \uparrow\uparrow \vec{a}$
3. For  $\lambda < 0$  is  $\lambda \cdot \vec{a} \uparrow\downarrow \vec{a}$
4. For  $\lambda = 0$  is  $\lambda \cdot \vec{a} = \vec{0}$

**Example:** On the whiteboard.

## Theorem

Let  $\vec{a}, \vec{b}$  be vectors and  $\lambda, \mu \in \mathbb{R}$ . Then  $\lambda \cdot \vec{a}$  has the following properties:

1.  $(\lambda + \mu) \cdot \vec{a} = \lambda \cdot \vec{a} + \mu \cdot \vec{a}$
2.  $(\lambda \cdot \mu) \cdot \vec{a} = \lambda \cdot (\mu \cdot \vec{a})$
3.  $\lambda \cdot (\vec{a} + \vec{b}) = \lambda \cdot \vec{a} + \lambda \cdot \vec{b}$

# Coordinate representation of vectors

## Remark

Even if no direction is required to describe something, it can still be useful to summarize numbers in a vector as the next example will show.

## Definition (Vector)

An  $n$ -dimensional vector  $\vec{x}$  is given by

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

with  $x_1 \in \mathbb{R}, \dots, x_n \in \mathbb{R}$ .

## Example: Coordinate representation of vectors in $\mathbb{R}^3$

To describe vectors, we consider a Cartesian coordinate system in the space through the origin  $O = (0, 0, 0)$  and the unit points  $E_1 = (1, 0, 0)$ ,  $E_2 = (0, 1, 0)$  and  $E_3 = (0, 0, 1)$  on the axes. The vectors determined by the unit points are called unit vectors:

$$\vec{e}_1 = \vec{OE}_1$$

$$\vec{e}_2 = \vec{OE}_2$$

$$\vec{e}_3 = \vec{OE}_3$$

## Remark

The vector  $\vec{a} = \vec{OA}$  to the point  $A = (a_1, a_2, a_3)$  can we then express as a unique linear combination of the unit vectors:

$$\vec{a} = a_1 \cdot \vec{e}_1 + a_2 \cdot \vec{e}_2 + a_3 \cdot \vec{e}_3$$

## Remark (Column notation)

$$\vec{a} = a_1 \cdot \vec{e}_1 + a_2 \cdot \vec{e}_2 + a_3 \cdot \vec{e}_3 = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

## Definition (Arithmetic operations with vectors)

Let

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ and } \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

### ► Addition and subtraction

We have

$$\vec{x} \pm \vec{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \pm \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 \pm y_1 \\ x_2 \pm y_2 \\ \vdots \\ x_n \pm y_n \end{pmatrix}$$

► Multiplication with a scalar

For  $\lambda \in \mathbb{R}$  we have

$$\lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot x_1 \\ \lambda \cdot x_2 \\ \vdots \\ \lambda \cdot x_n \end{pmatrix}$$

► Examples (3)

## Remark

As we combine different vectors with each other, often linear combinations are created.

## Definition (Linear combination)

Let  $\vec{x}_1, \dots, \vec{x}_m$  be  $n$ -dimensional vectors and  $c_1 \in \mathbb{R}, \dots, c_m \in \mathbb{R}$ .  
Then

$$c_1 \cdot \vec{x}_1 + c_2 \cdot \vec{x}_2 + \dots + c_m \cdot \vec{x}_m = \sum_{i=1}^m c_i \cdot \vec{x}_i$$

is a linear combination of the vectors  $\vec{x}_1, \dots, \vec{x}_m$ .

# Dot/Inner product

## Definition

1. In 2-dimensional space, let  $\vec{a} = (a_1, a_2)$  and  $\vec{b} = (b_1, b_2)$ . We define their dot product of  $\vec{a}$  and  $\vec{b}$  to be

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1 b_1 + a_2 b_2$$

2. In 3-dimensional space, their dot or inner product is defined to be

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1 b_1 + a_2 b_2 + a_3 b_3$$

3. Generally, in  $n$ -dimensional space, their dot or inner product is defined to be:

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

**Examples:** On the whiteboard.

**Remark**

We can only compute the dot product of two vectors if they have the same dimensions. For example, it is not possible to compute the dot product as follows:

$$\vec{x} \cdot \vec{y} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}.$$

## Theorem

Let  $\vec{a}, \vec{b}, \vec{c}$  be vectors and  $\lambda \in \mathbb{R}$ .

1.  $(\lambda \cdot \vec{a}) \cdot \vec{b} = \lambda \cdot (\vec{a} \cdot \vec{b})$
2.  $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$
3.  $\vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}$

- ▶ Examples (2)

## Definition (Orthogonal vectors)

Two vectors  $\vec{a}$  and  $\vec{b}$  are called orthogonal if  $\vec{a} \cdot \vec{b} = 0$ .

Examples: On the whiteboard.

# The norm or magnitude of a vector

## Definition

Let  $\vec{a}$  be a vector in  $\mathbb{R}^n$ . We define the norm or magnitude of  $\vec{a}$ , and denote by  $|\vec{a}|$ , the number

$$|\vec{a}| = \sqrt{\vec{a} \cdot \vec{a}}.$$

## Examples:

1. For

$$\vec{a} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix},$$

the norm of  $\vec{a}$  is

$$|\vec{a}| = \sqrt{\lambda_1^2 + \lambda_2^2}$$

2. For  $\vec{a} = (\lambda_1, \lambda_2, \lambda_3)$ , the norm of  $\vec{a}$  is

$$|\vec{a}| = \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}$$

# The norm and geometric length

**Notes:** When  $n = 2$  and  $n = 3$ , then the definition of norm is compatible with the geometric length by using Pythagoras theorem.

**Proof 1:** Applying the Theorem of Pythagoras

$$a^2 = \lambda_1^2 + \lambda_2^2$$

**Proof 2:**

$$\begin{aligned} |\vec{a}|^2 &= |\vec{OQ}|^2 + |\vec{QP}|^2 = |\vec{OR}|^2 + |\vec{RQ}|^2 + |\vec{QP}|^2 = \\ &\quad \lambda_1^2 + \lambda_2^2 + \lambda_3^2 \end{aligned}$$

**Definition (Distance between two points in the space)**

The distance between two points,  $A = (a_1, a_2, a_3)$  and  $B = (b_1, b_2, b_3)$ , in the space is given by:

$$d = |\vec{A} - \vec{B}| = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + (b_3 - a_3)^2}$$

- Example

## General Pythagoras theorem

### Theorem

Let  $\vec{a}$  and  $\vec{b}$  be vectors in  $\mathbb{R}^n$ . Then, vectors  $\vec{a}$  and  $\vec{b}$  are perpendicular if and only if

$$|\vec{a} + \vec{b}|^2 = |\vec{a}|^2 + |\vec{b}|^2.$$

# Angle between two vectors

## Remark

How can we calculate the angle  $\varphi$  between two vectors  $\vec{a}$  and  $\vec{b}$  in term of their coordinates?

- ▶ Example

### Theorem

Let  $\vec{a}$  and  $\vec{b}$  be vectors and  $\varphi$  the angle between these vectors.  
Then,

$$\vec{a} \cdot \vec{b} = |\vec{a}| \cdot |\vec{b}| \cdot \cos(\varphi)$$

- ▶ Example

### Remark

Let  $\vec{a}, \vec{b}$  be vectors and  $\varphi$  the angle between them. Then

$$\cos(\varphi) = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}| \cdot |\vec{b}|}$$

# Linear Algebra

Lectures 2: Linear equation systems

# Contents

- ▶ Linear equation systems (LES)
- ▶ Matrix representation of a LES
- ▶ Row-echelon form and Gaussian elimination
- ▶ Reduced row-echelon form and Gauss-Jordan elimination

# Linear equation systems

In general, if a linear relation exists between the inputs and outputs of a system, we call it a linear system.

## Remark (Applications of linear equation systems)

- ▶ Rotations in space
- ▶ Demand calculations for stock holding
- ▶ Modeling of customer streams
- ▶ Flow network
- ▶ Electrical circuits
- ▶ Quantum mechanics

## Definition (Linear equation)

A linear equation in  $n$  variables:

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

where

$a_1, \dots, a_n, b$  : real numbers

$x_1, \dots, x_n$  : unknown or variables

$a_1, \dots, a_n$  : are called coefficients of the equation.

$a_1$  : leading coefficient or pivot if non-zero

## Examples

- ▶  $2x + y - 4z = 0$

- ▶  $2x + y - 4z = 3$

## Definition (Linear equation system (LES))

A system of  $m$  linear equations is called a linear equation system:

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{array} \right. \quad (1)$$

where

$a_{ij}, b_i, i = 1, \dots, m, j = 1, \dots, n$  : real numbers;

$x_1, \dots, x_n$  unknowns or variables.

$a_{ij}$  coefficients.

## Examples

1. 
$$\begin{cases} x + y = 3 \\ x - y = -1 \end{cases}$$

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

3. 
$$\begin{cases} x + y = 3 \\ x + y = 1 \end{cases}$$

## Solution set of a LES

- ▶ A **solution** of the linear equation system (1) in  $n$  variables is a set of numbers  $s_1, s_2, \dots, s_n$  such that for all  $i = 1, 2, \dots, n$ , then

$$a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n = b_i.$$

- ▶ Solution set is the set of all solutions of (1).

## Consistency of a LES

A system of  $m$  linear equations in  $n$  variables:

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

- ▶ **consistent** if it has at least one solution;
- ▶ **inconsistent** if it has no solution;
- ▶ If  $b_i = 0$  for all  $i = 1, \dots, m$ , then the LES is **homogeneous**, otherwise **inhomogeneous**.

**Notes:** Every system of linear equations has either

1. exactly one solution,
2. infinitely many solution, or
3. no solution.

## Equivalent systems

- ▶ Two systems of linear equations are called **equivalent** if they have precisely the same solution set.
- ▶ **Notes:**  
**Eliminations:** We use following operations on a system of linear equations to produce an equivalent system:
  1. Interchange two equations
  2. Multiply an equation by a nonzero constant
  3. Add a multiple of an equation to another equation.

## Example

Solve a system of linear equations

$$\begin{cases} x - 2y + 3z = 9 \\ -x + 3y = -4 \\ 2x - 5y + 5z = 17 \end{cases} \quad (2)$$

**Solution:** Details on the whiteboard.

## Coefficient matrix of a LES I

- The coefficient matrix of the LES 1 is defined by

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},$$

- Notes:

1. Every entry  $a_{ij}$  in a matrix is a real number;
2. A matrix with  $m$  rows and  $n$  columns is said to be of size  $m \times n$ ;
3. If  $m = n$ , then the matrix is called square of order  $n$ ;
4. For a square matrix, the entries  $a_{11}, a_{22}, \dots, a_{nn}$  are called the main diagonal entries.

## Coefficient matrix of a LES II

- ▶ the right-hand side with the vector

$$\vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

- ▶ Matrix form of the LES 1:  $Ax = b$ , where

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

# Augmented matrix

- Augmented matrix for the LES (1)

$$[A|\vec{b}] = \left( \begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \cdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right)$$

- The  $n$  numbers  $x_1, \dots, x_n$  of the solution of the LES is called solution vector:

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$$

- Example: The augmented matrix of the LES (2) is

$$[A|b] = \left( \begin{array}{cccc|c} 1 & -2 & 3 & 9 & 9 \\ -1 & 3 & 0 & -4 & -4 \\ 2 & -5 & 5 & 17 & 12 \end{array} \right).$$

## Gaussian elimination: Row echelon form

- ▶ It is especially easy to solve a LES, if the coefficient matrix is given in row echelon form.
- ▶ A matrix is in row echelon form if the following three conditions are fulfilled:
  1. All nonzero rows are above any rows of all zeros.
  2. The leading entry (the pivot or the leftmost non-zero entry) in a nonzero row is 1.
  3. Each leading 1 of a row is in a column to the right of the leading 1 of the row above it.
- ▶ Note: All entries in a column below a leading entry are zeros.

## Examples I

- By elimination, the augmented matrix of the LES (2) could be transformed to the row-echelon form:

$$\left( \begin{array}{ccc|c} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{array} \right)$$

- The augmented matrix

$$\left( \begin{array}{ccccc|c} 0 & 2 & -3 & 4 & 1 & 7 \\ 0 & 0 & 0 & 5 & 2 & 4 \\ 0 & 0 & 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

represents the system of linear equations

## Examples II

$$\begin{cases} 2y - 3z + 4w + t = 7 \\ 5w + 2t = -4 \\ -3t = 1 \end{cases}$$

## Back-substitution for solving a LES with a coefficient matrix in row echelon form

1. We solve the equations for the leading variables (i.e. the variables corresponding to leading entries, also called **basic variables**).
2. Starting with the last (= bottom) we substitute the calculated variables into the equations above.
3. **Free variables** (i.e. variables not corresponding to leading entries) are assigned arbitrary values.

### Remark

If the rows without leading ones are contradictory, then the solution is the empty set. Hence, the LES is inconsistent.

- ▶ Examples (3)

### Remark

An arbitrary LES is usually not in row echelon form. We must then first transform the system to the row echelon form. The transformation must be done in such way that the solution set remains the same as for the original system.

### Remark

Two LES with the same set of solutions are called equivalent.

## Remark (Elementary row operations)

- ▶ **Interchange:** Interchange two rows:  $r_{ij} : R_i \leftrightarrow R_j$
- ▶ **Scaling:** Multiply all entries in a row by a nonzero constant:  
 $r_i^{(k)} : kR_i \rightarrow R_i$
- ▶ **Replacement:** Replace one row by the sum of itself and a multiple of another row:  $r_{ij}^{(k)} : kR_i + R_j \rightarrow R_j$ .

## Theorem

If the augmented matrix  $(A^* | \vec{b}^*)$  is transformed from  $(A | \vec{b})$  by the use of a finite number of elementary row operations, then the solution sets to these two systems are equivalent. We say that the systems are row equivalent.

### Remark (Gauss elimination)

A given augmented matrix is used as a starting point:  $(A|\vec{b})$ .

Using elementary row operations, this augmented matrix is transformed into an equivalent augmented matrix:  $(A^*|\vec{b}^*)$  with  $A^*$  in row echelon form.

We proceed as follows:

1. We look for the leftmost column with at least one nonzero element.
2. In this column, we look for the element with the largest absolute value. We interchange this row and the first row. (Pivotsearch to obtain numerical stability).
3. If the uppermost number  $a \neq 0$  is, then we multiply the first row by  $\frac{1}{a}$  to create a leading 1.
4. We add a multiple of the first row two the other rows, to create zeros below the leading 1.
5. These steps are repeated for a submatrix. The submatrix is obtained, by covering the first row (or the actual first row and all rows above it). The steps are repeated until the whole matrix is in row echelon form.

Examples:

$$1. \begin{bmatrix} 0 & 0 & -2 & 0 & 8 & 12 \\ 2 & 8 & -6 & 4 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & 4 \end{bmatrix}$$

$$2. \begin{bmatrix} 1 & -2 & 3 & 9 \\ -1 & 3 & 0 & -4 \\ 2 & -5 & 5 & 17 \end{bmatrix}$$

**Gauss-Jordan elimination:** The process for reducing a matrix to a reduced row-echelon form which satisfies the following:

1. it is a row-echelon form;
2. for each non-zero row, the first non-zero entry from the left is equal to 1 and it is called **pivot** or **leading 1**.
3. Every column that has a pivot/leading 1 has zeros in every position above and below its pivot.

**Example:** On the whiteboard.

## Definition (Rank)

Given a matrix  $A$ . The number of the nonzero rows of  $A^*$ , is called the rank of matrix  $A$  and denoted by  $r(A)$  or simply  $r$  in case  $A$  is shown.

## Remark (Solutions to a LES)

Using the augmented matrix  $(A^* | \vec{b}^*)$  in row echelon form, we can identify the kind of solutions that the corresponding LES has (no solution, exactly one solution, infinitely many solutions).

Back-substitution can be used to determine the set of solutions.

## Theorem (The existence of solutions)

The LES with the augmented matrix  $(A|\vec{b})$  is consistent if after applying the Gaussian elimination algorithm we arrive at the augmented matrix  $(A^*|\vec{b}^*)$  with:

$$b_{r+1}^* = \dots = b_m^* = 0$$

where  $A$  is an  $m \times n$  matrix.

Then we have:

- ▶  $r = n \Rightarrow$  the LES has exactly one solution
- ▶  $r < n \Rightarrow$  the LES has infinitely many solutions (and  $n - r$  of the  $x_j$ :s can be taken arbitrary)
  
- ▶ Examples (5)

### Remark (Homogeneous linear equation systems)

A LES with augmented matrix  $(A|\vec{0})$  is called a **homogeneous LES**. The right-hand side of a homogeneous LES is always the null vector and has at least the trivial solution  $\vec{x} = \vec{0}$ .

## Theorem

For a homogeneous LES, we have:

- ▶  $r = n \Leftrightarrow$  the LES has only the trivial solution  $\vec{0}$
- ▶  $r < n \Leftrightarrow$  the LES has infinitely many solutions and  $n - r$  of the  $x_j$ :s can be taken arbitrary.

## Remark

A homogeneous LES with  $n > m$  (with more unknowns than equations) has infinitely many solutions.

### Remark

For a homogeneous LES is it sufficient to transform the coefficient matrix  $A$  with elementary row operations, since the null vector remains unchanged.

### Definition (Square linear equation systems)

A LES with  $m = n$  (Number of equations = Number of unknowns) is called a square LES.

## Theorem

*A square LES has a unique solution if and only if  $r = n$ .*

Remark (Solution of a square LES with a unique solution by Gauss-Jordan elimination)

We can obtain the solution in the following way from the augmented matrix  $(A^* | \vec{b}^*)$ :

1. We add a multiple of the last row to the other rows, to create zeros above the leading 1.
2. We repeat the first step for the submatrix obtained by covering the last row and repeat this until the first row.

The result is an augmented matrix of the form

$$(A^{**} | \vec{b}^{**}) = \left( \begin{array}{cccc|c} 1 & 0 & \cdots & 0 & b_1^{**} \\ 0 & 1 & \cdots & 0 & b_2^{**} \\ \vdots & & & \vdots & \vdots \\ 0 & \cdots & & 1 & b_n^{**} \end{array} \right)$$

i.e. we can immediately see the solution:

$$\vec{x} = \begin{pmatrix} b_1^{**} \\ b_2^{**} \\ \vdots \\ b_n^{**} \end{pmatrix}$$

### Remark

The Gaussian algorithm together with back-substitution is also called **Gauss-Jordan elimination**.

- ▶ Example (2)

# Linear Algebra

Lectures 3+4: Matrices

# Outline

- ▶ Introduction
- ▶ Matrix operations
- ▶ Special matrices
- ▶ The inverse of a matrix
- ▶ Matrix and linear equation system

## Introduction

1. Linear transformations like transition, reflection, rotation, can be considered as matrix actions.
2. Our ability to analyze and solve equations will be greatly enhanced when we can perform algebraic operations with matrices.
3. Transformations on equations of a linear system can be written as transformations on rows of a matrix.
4. Matrix algebra provides tools for manipulating matrix equations and creating various useful formulas in ways similar to doing ordinary algebra with real numbers.
5. The inverse of a matrix, if it exists, allows us to treat the matrix as a number which helps to find out the solution of the linear system explicitly.
6. Matrix algebra can be applied in many fields like economics, computer graphics, image processing, etc.

# Introduction

## Remark

A matrix is a rectangular array, filled with numbers.

- ▶ Example

## Definition (Matrix)

A rectangular array  $A$  of  $m \cdot n$  numbers in  $m$  rows and  $n$  columns

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

is called a **matrix of size  $m \times n$** .

$m$  is the number of rows of  $A$ ;

$n$  is the number of columns of  $A$ ;

$a_{ij}$ , for  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, n$ , are called entries of  $A$ .

$M_{m \times n}$ : the set of all matrices of size  $m \times n$ .

- ▶ Example:

$$\begin{pmatrix} 0 & 1 & -1 \\ 1 & -2 & 3 \end{pmatrix} \in M_{2 \times 3}, \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & -1 \\ -1 & 3 \end{pmatrix} \in M_{3 \times 2}$$

### Remark

A vector is only a special case of a matrix.

### Definition

A matrix consisting of just one column is called a column vector:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$$

## Definition

A matrix consisting of just one row is called a **row vector**:

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$$

## Definition (Equality of matrices)

Two matrices  $A$  and  $B$  are equal, if they have the same size  $m \times n$  and if all elements are equal:

$$a_{ij} = b_{ij} \text{ for all } i = 1, \dots, m, j = 1, \dots, n$$

# Matrix Operations: Addition and Subtraction of Matrices

## Definition (Addition and Subtraction of Matrices)

Let

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

and

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

be  $m \times n$  matrices.

Then

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

i.e. the matrices are added and subtracted element-wise.

- ▶ Example: On the whiteboard

## Matrix Operations: Multiplication with a scalar

- ▶ Example: On the whiteboard

### Definition (Multiplication with a Scalar)

Matrices are element-wise multiplied by a scalar. Let

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

be an  $m \times n$  matrix and let  $\lambda \in \mathbb{R}$ .

Then

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

- ▶ Example: On the whiteboard

## Theorem (Calculation Rules)

Let  $A, B, C$  be  $m \times n$  matrices and  $\lambda, \mu \in \mathbb{R}$

1.  $A + B = B + A$  (Commutative)
2.  $(A + B) + C = A + (B + C)$  (Associative)
3.  $\lambda \cdot (A + B) = \lambda \cdot A + \lambda \cdot B$  (Distributive)
4.  $(\lambda + \mu) \cdot A = \lambda \cdot A + \mu \cdot A$  (Distributive)

- ▶ Example: On the whiteboard

## Matrix Operations: Matrix Multiplication

- ▶ Example

### Definition (Matrix multiplication)

For an  $m \times n$  matrix  $A$  and an  $n \times p$  matrix  $B$ . The matrix  $m \times p$  matrix  $C$  with the elements

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$$

called the matrix product of  $A$  and  $B$ , denoted by  $C = AB$ .

## Remark

- ▶ The matrix product  $AB$  is only defined if the number of columns of  $A$  is equal to the number of rows of  $B$ .
- ▶ Write the sizes of  $A$  and  $B$  beside each other:

$$m \times n \text{ and } q \times p$$

The matrix product is only defined if the inner numbers are equal, i.e.  $n = q$ . Then the outer numbers tell the size of the matrix product:  $m \times p$ .

## Remark

To determine the element in the  $i$ -th row and the  $j$ -th column of  $AB$ , we multiply the elements of the  $i$ -th row of  $A$  element-wise with the elements of the  $j$ -th column of  $B$  and add the generated products:

$$\begin{pmatrix} & \vdots & \\ a_{i1} & \cdots & a_{in} \\ & \vdots & \end{pmatrix} \cdot \begin{pmatrix} & b_{1j} & \\ \dots & \vdots & \dots \\ & b_{nj} & \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \cdots & c_{ij} & \cdots \\ & \vdots & \end{pmatrix}$$

- ▶ Examples (2): On the whiteboard.

## Remark

The product of an  $m \times 1$  column vector with a  $1 \times m$  row vector is an  $m \times m$  matrix:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \cdot \begin{pmatrix} b_1 & \cdots & b_m \end{pmatrix} = \begin{pmatrix} a_1 b_1 & \cdots & a_1 b_m \\ \vdots & & \vdots \\ a_m b_1 & \cdots & a_m b_m \end{pmatrix}$$

## Remark

The product of a  $1 \times m$  row vector with an  $m \times 1$  column vector is a  $1 \times 1$  matrix, i.e. a scalar (only a number):

$$\begin{pmatrix} b_1 & \cdots & b_m \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} = a_1 b_1 + \cdots + a_m b_m$$

This is a dot product written with matrix notation!

- ▶ Example

## Remark (Missing Calculation Rules)

Many calculation rules for scalars hold also for matrices. But for the multiplication of matrices the commutative law does not hold. In general, is the following NOT fulfilled:

- ▶ Commutative law:  $AB = BA$
- ▶ Zero law for products:  $AB = 0 \Rightarrow A = 0$  or  $B = 0$ .
- ▶ Cancellation law:  $AC = BC$  and  $C \neq 0 \Rightarrow A = B$

## Remark (Missing commutative law)

There are some situations, where the commutative law is missing for matrix multiplications:

- ▶  $AB$  is defined, but  $BA$  is not defined.
- ▶  $AB$  and  $BA$  are of different sizes.
- ▶ It is actually also possible that  $AB \neq BA$  is not fulfilled, even if  $AB$  and  $BA$  have the same size:

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

but

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

## Matrix Operations: Transposing Matrices

### Definition (Transposed matrix)

The  $n \times m$  matrix that we obtain if we write the rows of the  $m \times n$  matrix  $A$  as columns, is called **the transposed matrix** of  $A$ ,

$$A^T = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \cdots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix}$$

- ▶ Example: On the white board.

# Properties of transposes

## Theorem

Let  $A$  and  $B$  be matrices. Then

1.  $(A^T)^T = A$
2.  $(A + B)^T = A^T + B^T$
3.  $(cA)^T = cA^T$
4.  $(AB)^T = B^T A^T$

# Some Special Matrices

Definition (Zero matrix)

$a_{ij} = 0$  for all  $i, j$ , i.e.

- ▶ Example:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

## Properties of zero matrices

### Theorem

Let  $A$  be an  $m \times n$  matrix and let  $c$  be a scalar. Then

1.  $A + 0_{m \times n} = A$
2.  $A + (-A) = 0_{m \times n}$
3. if  $cA = 0_{m \times n}$ , then either  $c = 0$  or  $A = 0_{m \times n}$

## Definition (Square matrix)

An  $n \times n$  matrix and the elements  $a_{ii}$ ,  $i = 1, \dots, n$  constitute the main diagonal.

- ▶ Example

## Definition (Diagonal matrix)

A square matrix

$$d_{ij} = 0 \text{ for } i \neq j,$$

i.e.

$$D = \begin{pmatrix} d_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_{nn} \end{pmatrix}$$

- ▶ Example

## Definition (Identity matrix)

A diagonal matrix with main diagonal elements = 1

$$d_{ij} = 0 \text{ for } i \neq j,$$

and

$$d_{ij} = 1 \text{ for } i = j,$$

is called the identity matrix and denoted by  $I_{n \times n}$  or  $I$  in case  $n$  is shown.

- ▶ Example:  $I_{2 \times 2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ , and  $I_{3 \times 3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ ,

## Definition (Upper triangular matrix)

A square matrix with

$$d_{ij} = 0 \text{ for } i > j,$$

i.e.

$$D = \begin{pmatrix} d_{11} & \cdots & d_{1n} \\ \vdots & & \vdots \\ 0 & \cdots & d_{nn} \end{pmatrix}$$

- ▶ Example: On the whiteboard

## Definition (Lower triangular matrix)

A square matrix with

$$d_{ij} = 0 \text{ for } i < j,$$

i.e.

$$D = \begin{pmatrix} d_{11} & \cdots & 0 \\ \vdots & & \vdots \\ d_{n1} & \cdots & d_{nn} \end{pmatrix}$$

- ▶ Example: On the whiteboard

## Definition (Symmetric matrix)

A square matrix  $A$  with

$$a_{ij} = a_{ji} \text{ for all } i, j,$$

is called symmetric. In other words,  $A$  is symmetric if and only if  $A = A^T$ .

- ▶ Examples: Transportation matrices, social network matrices, image processing matrices...
- ▶  $\begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & -2 \\ -1 & -2 & 3 \end{pmatrix}$  is symmetric and  $\begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & \textcircled{-2} \\ -1 & \textcircled{0} & 3 \end{pmatrix}$  is not symmetric

# The inverse of a matrix

## Definition (Regular matrices)

An  $n \times n$  matrix  $A$  is called **regular** if  $\text{rank}(A) = n$  and **singular** if  $\text{rank}(A) < n$ .

## Remark

$$A \text{ regular} \Leftrightarrow$$

$$\text{rank}(A) = n \Leftrightarrow$$

Every LES  $A\vec{x} = \vec{b}$  has exactly one solution  $\Leftrightarrow$

The homogeneous LES  $A\vec{x} = \vec{0}$  has only the trivial solution

## Theorem

If the  $n \times n$  matrix  $A$  is regular, then a uniquely determined matrix  $n \times n$  matrix  $X$  with  $AX = I$  exists.

## Definition (Invertible matrix)

An  $n \times n$  matrix  $A$  is called invertible if an  $n \times n$  matrix  $X$  exists with  $AX = XA = I$ , where  $I$  is the identity matrix of size  $n \times n$ .

## Remark

The matrix  $X$  with  $AX = I$  is called **the inverse** to  $A$ , and is written  $X = A^{-1}$ .

## Remark

1.  $A$  regular  $\Leftrightarrow A$  invertible.
2. If  $A$  is invertible, then is  $X$  uniquely determined.
3. For the inverse  $A^{-1}$ , the following holds:  $A^{-1}$  is regular and

$$AA^{-1} = A^{-1}A = I$$

with  $I$  the identity matrix.

## Power of a square matrix

Given a square matrix  $A \in M_{n \times n}$ . We define

1.  $A^0 = I$

2.  $A^k = \underbrace{AA \cdots A}_{k \text{ times}} (k \in \mathbb{N})$ .

3.  $A^r \cdot A^s = A^{r+s}$

4.  $(A^r)^s = A^{rs}$

5. If  $D = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}$ , then  $D^k = \begin{pmatrix} d_1^k & 0 & \dots & 0 \\ 0 & d_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n^k \end{pmatrix}$

## Theorem (Properties of inverse matrices)

Let  $A$  be an invertible matrix, let  $k$  be a positive integer, and  $c$  be a scalar not equal to zero. Then

1.  $A^{-1}$  is invertible and  $(A^{-1})^{-1} = A$
2.  $A^k$  is invertible and  $(A^k)^{-1} = (A^{-1})^k$
3.  $cA$  is invertible and  $(cA)^{-1} = \frac{1}{c}A^{-1}$
4.  $A^T$  is invertible and  $(A^T)^{-1} = (A^{-1})^T$
5. If  $A$  and  $B$  are invertible, then  $AB$  is invertible and  $(AB)^{-1} = B^{-1}A^{-1}$ .

## Computation of the inverse

1. Augment  $A$  to the right with the identity matrix
2. Perform a Gauss-Jordan elimination

Then we obtain the inverse matrix  $A^{-1}$  on the right-hand side:

$$(A|I) = (I|A^{-1})$$

- ▶ Notes:
  1. If  $A$  cannot be row reduced to  $I$ , then  $A$  is singular.
  2. At first is it usually not obvious if the matrix  $A$  is invertible at all. If this is not the case, then we would arrive at a zero row at the left-hand side of the augmented matrix. We would then conclude that  $A$  is singular and stop the process.

## Examples

- ▶ Find the inverse of the following matrix  $\begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ -6 & 2 & 3 \end{pmatrix}$
- ▶ Solution: On the whiteboard.

## Notes

From the definitions for matrix operations, a linear system of  $m$  equations and  $n$  variables  $x_1, \dots, x_n$  can be written as matrix operation:

$$Ax = b,$$

where  $A$  is the coefficient matrix of the system,  $x$  (resp.  $b$ ) represents column vector of variables (the right hand side) of the system.

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

## Theorem

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

Then the LES  $A\vec{x} = \vec{b}$  has the unique solution

$$\vec{x} = A^{-1}\vec{b}$$

### ► Example

## Remark

We can use this theorem if we have to solve many LES with the same coefficient matrix  $A$  but different right-hand sides  $\vec{b}$ .

# Linear Algebra

Lectures 5+6: Determinants and applications

## Outlines

- ▶ Determinant of a matrix
- ▶ Properties of determinants
- ▶ Gaussian elimination for calculating determinants
- ▶ Applications in inverse finding and solution solving

## Determinants

Determinant is a number assigned to each square matrix. A single number, determinant, can tell only so much about a matrix. Still, it is amazing how much this number can do. We can use determinant for

1. the invertibility existence of a matrix;
2. the solution existence of a linear system;
3. finding the inverse of a matrix using cofactors;
4. solving explicitly a linear system by Cramer's rule;
5. measuring the dependence of  $A^{-1}b$  on each element of  $b$ . If one parameter is changed in an experiment, or one observation is corrected, the "influence coefficient" in  $A^{-1}$  is a ratio of determinants.
6. measuring the amount by which a linear transformation changes the area of a figure. Finding the volume of a box in  $n$ -dimensional space.

## Determinant of a $2 \times 2$ matrix

The number

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

is called **the determinant** of the  $2 \times 2$  matrix

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

## Determinant of a $3 \times 3$ matrix

The number

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$
$$= a_{11} \cdot M_{11} - a_{12} \cdot M_{12} + a_{13} \cdot M_{13}$$

is called **the determinant** of the  $3 \times 3$  matrix

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

where  $M_{ij}$  is **the determinant of the matrix obtained from  $A$  by removing its row  $i$  and column  $j$** , and called **minor** of the entry  $a_{ij}$ .

## Remarks

- ▶ For the computation of the determinant of a  $3 \times 3$  matrix, we can use [Sarrus' rule](#). Accordingly, its determinant is equal to the difference of the down-diagonal sum and up-diagonal sum of the following matrix:

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

- ▶ We could also compute the determinant by expanding one row or one column and then compute the determinant for the remaining  $2 \times 2$  matrices. This can also generalized and then applied to the determinant for an  $n \times n$  matrix.

## Cofactor expansion

In general, the determinant of a square matrix of size  $n \times n$  is a number defined inductively as follow:

1. Pick any one row (column) of  $A$ ;
2. For each entry in the row chosen, find its co-factor;
3. Multiply each entry in the row (column) chosen by its co-factor and take the sum that results the determinant of  $A$ .

## Determinant of an $n \times n$ matrix

- ▶ Minor  $M_{ij}$  of the entry  $a_{ij}$ : the determinant of the  $(n - 1) \times (n - 1)$  matrix obtained from  $A$  by removing its row  $i$  and column  $j$ .
- ▶ Cofactor of  $a_{ij}$

$$C_{ij} = (-1)^{i+j} M_{ij}.$$

- ▶ The determinant of  $A$  is given by
  1. Cofactor expansion along the row  $i$ :

$$\det A = |A| = \sum_{j=1}^n a_{ij} C_{ij} = a_{i1} C_{i1} + a_{i2} C_{i2} + \cdots + a_{in} C_{in}.$$

2. Cofactor expansion along the column  $j$ :

$$\det A = |A| = \sum_{i=1}^n a_{ij} C_{ij} = a_{1j} C_{i1} + a_{2j} C_{i2} + \cdots + a_{nj} C_{nj}.$$

## Examples

- ▶ **Notes:** The row (or column) containing the **most zeros** often if the best choice for expansion by cofactors.
- ▶ **Examples:** If  $A$  is an **triangular matrix**, then its determinant is the product of the entries on the main diagonal, where
  1. **upper triangular matrix:** All the entries below the main diagonal are zeros;
  2. **lower triangular matrix:** All the entries above the main diagonal are zeros;
  3. **diagonal matrix:** All the entries above and below the main diagonal are zeros.
- ▶ **Example:** Find the determinant

$$\begin{vmatrix} 2 & 1 & 0 \\ 1 & 1 & 4 \\ -3 & 2 & 5 \end{vmatrix}$$

**Solution:** On the white board.

## Numerical note

By today's standard, a  $25 \times 25$  matrix is small. Yet it would be impossible to calculate a  $25 \times 25$  determinant by cofactor expansion. In general, a cofactor expansion requires over  $n!$  multiplications, and  $25!$  is approximately  $1.5 \times 10^{25}$ , it would spend approximately 500,000 years to compute a  $25 \times 25$  determinant by this method.

# Basic properties

## Theorem

- ▶ If  $A$  contains a zero row or a zero column, then we have  $\det(A) = 0$ .
- ▶  $\det(A^T) = \det(A)$
- ▶ If  $A$  is a triangular matrix, then we have:

$$\det(A) = a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn}$$

- ▶  $\det(I) = 1$ , where  $I$  is the identity matrix.
- ▶  $\det(AB) = \det(A)\det(B)$ . Consequently,  $\det(A^{-1}) = \frac{1}{\det(A)}$ .

**Remark:** In general,  $\det(A + B) \neq \det(A) + \det(B)$ .

## Row operation rules for determinants

Theorem (Determinants and elementary row operations)

Let  $A$  be a square matrix

1. If a multiple of one row (column) of  $A$  is added to another row (column) to produce a matrix  $B$ , then  $\det B = \det A$ .
2. If two rows of  $A$  are interchanged to produce  $B$ , then  $\det B = -\det A$
3. If one row of  $A$  is multiplied by  $k$  to produce  $B$ , then  $\det B = k \det A$ .

Proof.

On the white board.



**Remark:** Most computer programs that compute  $\det A$  for a general matrix  $A$  use the elementary row operations to reduce  $A$  to triangular matrix.

## Examples

Compute  $\det(A)$ , where  $A = \begin{pmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{pmatrix}$

**Solution:** On the whiteboard!

## Zero determinants

If  $A$  is a square matrix and any of the following conditions is true, then  $\det(A) = 0$ .

1. An entire row (column) consists of zeros;
2. Two rows (columns) are equal;
3. One row (resp. column) is a multiple of another row (resp. column).
4. One row (resp. column) is a linear combination of other rows (resp. column).

## Examples

$$1. \begin{vmatrix} 1 & k & k^2 \\ 1 & k & k^2 \\ 1 & k & k^2 \end{vmatrix} = 0$$

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

$$3. \begin{vmatrix} a_1 & b_1 & a_1 + b_1 \\ a_2 & b_2 & a_2 + b_2 \\ a_3 & b_3 & a_3 + b_3 \end{vmatrix} = 0$$

## Theorem

For a square matrix  $A$  we have:

$A$  is regular  $\Leftrightarrow$

$A$  is invertible  $\Leftrightarrow$

$$\det(A) \neq 0$$

## Remark

If  $A$  invertible, then

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

## Theorem (Cramer's rule)

If  $A$  is an invertible square matrix (i.e.  $\det(A) \neq 0$ ), then the LES  $A\vec{x} = \vec{b}$  has the unique solution

$$\vec{x} = \frac{1}{\det(A)} \begin{pmatrix} D_1 \\ \vdots \\ D_n \end{pmatrix}$$

Here  $D_k$  is the determinant we would obtain if in  $\det(A)$  the  $k$ -th column is replaced by  $\vec{b}$ .

Proof.

On the whiteboard. □

## Remarks:

- ▶ Using determinants we can give a very nice and explicit formula for the solution to a linear system with  $n$  variables and  $n$  equations in case it has unique solution.
- ▶ Cramer's rule is needed in a variety of theoretical calculations. However, the formula is inefficient for hand calculations, except for  $2 \times 2$  or perhaps  $3 \times 3$  matrices because of the determinant computing complexity.
- ▶ In general, for solving a linear system, Gaussian and Gaussian-Jordan eliminations are used more often.

## Examples

Determine the value of  $s$  for which the system has a unique solution, and use Cramer's rule to describe the solution

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

### Solution:

- View the system as  $Ax = b$ . Then

$$A = \begin{bmatrix} 3c & -2 \\ -6 & c \end{bmatrix}, \quad D_1 = \begin{bmatrix} 4 & -2 \\ 1 & c \end{bmatrix}, \quad D_2 = \begin{bmatrix} 3c & 4 \\ -6 & 1 \end{bmatrix},$$

- The system has unique solution if and only if  $\det(A) \neq 0$ :

$$\det(A) = 3c^2 - 12 = 3(c - 2)(c + 2) \neq 0 \Leftrightarrow c \neq \pm 2.$$

- When  $c \neq \pm 2$ , the solution to system is

$$x = \frac{\det(D_1)}{\det A} = \frac{4c + 2}{3(c - 2)(c + 2)}.$$

$$y = \frac{\det(D_2)}{\det A} = \frac{c + 8}{(c - 2)(c + 2)}.$$

## Inverse matrix finding: Case of $2 \times 2$ matrix

Recall:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

# A formula for the inverse matrix

## Theorem

Let  $A$  be an invertible matrix and let  $C$  be cofactor matrix of  $A$ , i.e.  $C_{ij}$  are cofactors of entries  $a_{ij}$ . Then,

$$A^{-1} = \frac{1}{\det A} C^T.$$

**Remark:** The matrix  $C^T$  is called the adjoint matrix of  $A$ , denoted by  $\text{adj}(A)$ .

**Proof.**

On the white board. □

## Examples

Using cofactors for finding the inverse of

1.

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

2.

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

**Solution:** On the white board.

# Linear Algebra

Lectures 7+8+9+10+11: Vector spaces

# Introduction

- ▶ Actually, a study of vector spaces is not much different from a study of  $\mathbb{R}^n$  itself. We can use our geometric experience with  $\mathbb{R}^2$ ,  $\mathbb{R}^3$  to visualize many general concepts.
- ▶ We can use vector space terminology to tie together important facts about rectangular matrices like rank concept.
- ▶ In applications of linear algebra, subspaces of  $\mathbb{R}^n$  usually arise in one of two ways: as the set of solutions of a **homogeneous linear system**, or as **the set of all linear combinations** of certain specified vectors. These lead to considering the **null** and **column spaces** of a matrix/linear transformation.

# Outline

- ▶ Vector Spaces
- ▶ Subspaces of Vector Spaces
- ▶ Spanning Sets and Linear Independence
- ▶ Basis and Dimension
- ▶ Rank of a Matrix and Systems of Linear Equations
- ▶ Four fundamental subspaces
- ▶ Linear transformations
- ▶ Coordinate systems

## Vector space I

Let  $V$  be a set on which two operations (**vector addition**:  $+ : V \times V \rightarrow V$  and **scalar multiplication**:  $\cdot : \mathbb{R} \times V \rightarrow V$ ) are defined. If the following axioms are satisfied for every  $u$ ,  $v$ , and  $w$  in  $V$  and every scalar (real number)  $\lambda$  and  $\mu$ , then  $V$  is called a *vector space*.

## Vector space II

1.  $u + v \in V$  (closed with the addition)
2.  $u + v = v + u$  (commutative addition)
3.  $u + v + w = u + (v + w)$  (associative addition)
4.  $V$  has a zero vector  $0$  s.t.  $u + 0 = u$  for all  $u \in V$
5. For each  $u$  in  $V$ , there is an opposite vector in  $V$ , denoted by  $-u$ , s.t.  $u + (-u) = 0$
6.  $\lambda u \in V$  (closed with scalar multiplication)
7.  $\lambda(u + v) = \lambda u + \lambda v$
8.  $(\lambda + \mu)u = \lambda u + \mu u$
9.  $\lambda\mu u = \lambda(\mu u)$
10.  $1.u = u$ .

## Examples I

1.  **$n$ -tuple space:**  $\mathbb{R}^n$ , with the vector addition and the multiplication with scalar.

► Vector addition

$$(u_1, u_2, \dots, u_n) + (v_1, v_2, \dots, v_n) = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n).$$

► Scalar multiplication

$$k(u_1, u_2, \dots, u_n) = (ku_1, ku_2, \dots, ku_n).$$

2. **Matrix space:**  $V = M_{m \times n}$  with the matrix addition and scalar multiplication. Example:  $m = n = 2$

► Matrix addition

$$\begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} + \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} = \begin{pmatrix} u_{11} + v_{11} & u_{12} + v_{12} \\ u_{21} + v_{21} & u_{22} + v_{22} \end{pmatrix}$$

► Scalar multiplication

$$k \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} = \begin{pmatrix} ku_{11} & ku_{12} \\ ku_{21} & ku_{22} \end{pmatrix}$$

## Examples II

3. The set of all real polynomials of degree not exceeding  $n$ :  
 $V = P_n(x)$  together with the polynomial addition and polynomial scalar multiplication forms a vector space. (verify this please!)
4. The set of all solutions of a homogeneous linear equation system together with vector addition and scalar multiplication forms a vector space. For instance, the solutions of the equation  $x + 2y - 4z = 0$ . (verify this please!)
5. The set of all continuous functions on a given domain with the scalar multiplication and function addition.

**Notes:** To show that a set is not a vector space, you need only find one axiom that is not satisfied.

**Examples:**

- ▶ The set of all integers is not a vector space since let  $\frac{1}{2} \in \mathbb{R}$  and  $1 \in V$  then  $(\frac{1}{2})(1) = \frac{1}{2} \notin V$ .
- ▶ The set of all second-degree polynomials is not a vector space, since let  $p(x) = x^2 - 2x - 1$  and  $q(x) = -x^2 - 1$ , then  $p(x) + q(x) = -2x - 2$  which is not a second-degree polynomial.
- ▶ The set of all solutions of a non-homogeneous linear equation system (verify this please).

# Subspaces

## Definition

$V$ : vector space

$W \subseteq V$  and  $W \neq \emptyset$ .

$W$  is called a **subspace** of  $V$  if  $W$  together with the addition and the multiplication with a scalar inherited from  $V$  is a vector space.

## Notes:

1. Two trivial subspaces of  $V$ : Zero vector space  $\{\vec{0}\}$ , and  $V$ .
2. Test for a subspace: a non-empty set  $W$  is a subspace of  $V$  if and only if it is closed with the addition and the multiplication with a scalar, i.e.,
  - If  $u$  and  $v$  are in  $W$ , then  $u+v$  is in  $W$
  - If  $u$  is in  $W$  and  $\lambda$  is any scalar, then  $\lambda u$  is in  $W$ .

## Examples

- ▶ A set of points on a line through the origin in the plane is a subspace of  $\mathbb{R}^2$ .
- ▶ Let  $W$  be the set of all  $2 \times 2$  symmetric matrices. Then  $W$  is a subspace of  $M_{2 \times 2}$ .
- ▶ The set of singular matrices of size  $2 \times 2$  is not a subspace of  $M_{2 \times 2}$ .
- ▶ The set of invertible matrices of size  $2 \times 2$  is not a subspace of  $M_{2 \times 2}$ .
- ▶ The set of solutions of a homogeneous system with  $n$  variables is a subspace of  $\mathbb{R}^n$ .

## Lemma

*The intersection of two subspaces is a subspace.*

# Linear combinations I

## Definition (Linear combinations)

Let  $V$  be a vector space and let  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  be vectors in  $V$ . Let  $\alpha_1, \alpha_2, \dots, \alpha_k$  be numbers. An expression of type

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k$$

is called a **linear combination** of  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ . The numbers  $\alpha_1, \alpha_2, \dots, \alpha_n$  are called the **coefficients** of the linear combination.

## Lemma

*The set of all linear combinations of vectors  $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$  in a vector space  $V$  is a subspace of  $V$ . This subspace is denoted by  $\text{span}(S)$  and called **the span** of  $S$ .*

## Examples

1. All linear combinations of vector  $(1, 2)$  in the plane  $\mathbb{R}^2$  are on the line through two points  $(1, 2)$  and the origin.
2. All linear combinations of vectors  $(1, 1, 0)$  and  $(1, 0, 0)$  in  $\mathbb{R}^3$  are on the plane through 3 points  $(0, 1, 0)$ ,  $(1, 0, 0)$  and the origin.

## Spanning set

1. The set of all linear combinations of  $k$  vectors  $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$  is called **the span of  $S$**  and denoted by  $\text{span}(S)$ . Hence,  
$$\text{span}(S) = \{a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k : \forall a_1, a_2, \dots, a_k \in \mathbb{R}\}.$$
2. A **spanning set of a vector space**: Given a vector space  $V$  and a set  $S \subseteq V$ . If every vector of  $V$  can be written as a linear combination of vectors in  $S$ , i.e.,  $V = \text{span}(S)$ , then  $S$  is called a **spanning set of  $V$** .

1. Some terminologies: Given a set of vectors  $S$ . If  $\text{span}(S) = V$ , then we can say that

- ▶  $S$  spans (generates)  $V$ ;
- ▶  $V$  is spanned by  $S$ ;
- ▶  $S$  is a spanning set of  $V$ .

2. Notes:

- i)  $\text{span}(\emptyset) = \{\vec{0}\}$ ;
- ii)  $S \subseteq \text{span}(S)$
- iii)  $S_1, S_2 \subseteq V$ , and  $S_1 \subseteq S_2$ , then  $\text{span}(S_1) \subseteq \text{span}(S_2)$ .

### Lemma

Let  $S = \{v_1, v_2, \dots, v_k\}$  be vectors in  $V$ . Then,

- i)  $\text{span}(S)$  is a subspace of  $V$ .
- ii)  $\text{span}(S)$  is the smallest subspace of  $V$  that contains  $S$ .

## Examples

1.  $e_1 = (1, 0, 0)$ ,  $e_2 = (0, 1, 0)$ ,  $e_3 = (0, 0, 1)$  is a spanning set for  $\mathbb{R}^3$ ;
2.  $e_1 = 1$ ,  $e_2 = x$ ,  $e_3 = x^2$  is a spanning set for  $P_2(x)$ ;
3. Let  $H$  be the set of all vectors of the form  $(a - 3b, b - a, a, b)$ , where  $a$  and  $b$  are arbitrary scalars. We can write down column vector

$$\begin{bmatrix} a - 3b \\ b - a \\ a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

which shows that  $H = \text{span}\{\mathbf{v}_1, \mathbf{v}_2\}$ , where  $\mathbf{v}_1 = (1, -1, 1, 0)$  and  $\mathbf{v}_2 = (-3, 1, 0, 1)$ .

Details are on the whiteboard!

# Linear independence and linear dependence

## Definition

Let  $S = \{v_1, v_2, \dots, v_k\}$  be a set of vectors in a vector space  $V$ .

Consider the equation

$$c_1 v_1 + \cdots + c_k v_k = \vec{0}.$$

- i) If the equation has only the trivial solution  
 $c_1 = c_2 = \cdots = c_k = 0$ , then  $S$  is called **linearly independent**.
- ii) If the equation has a nontrivial solution (i.e., not all zeros),  
then  $S$  is called **linearly dependent**.

# Notes

1.  $\emptyset$  is linearly independent
2. If  $\vec{0} \in S$ , then  $S$  is linearly dependent.
3. If  $v \neq \vec{0}$ , then  $\{v\}$  is linearly independent
4. If  $S_1 \subseteq S_2$ , then
  - ▶  $S_1$  is linearly dependent  $\Rightarrow S_2$  is linearly dependent;
  - ▶  $S_2$  is linearly independent  $\Rightarrow S_1$  is linearly independent;

## Examples: Testing for linearly independent

1. Vectors  $(1, 1)$  and  $(-3, 2)$  are linearly independent in  $\mathbb{R}^2$ .
2. Determine whether the following set of vectors in  $P_2$  is linearly independent or dependent.

$$S = \{1 + x - 2x^2, 2 + 5x - x^2, x + x^2\}.$$

3. Determine whether the following set of vectors in  $2 \times 2$  matrix space is linearly independent or dependent.

$$S = \left\{ \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \right\}$$

4. Let  $p_1(t) = 1$ ,  $p_2(t) = t$  and  $p_3(t) = 4 - t$  be polynomials.  
Then the set  $\{p_1, p_2, p_3\}$  are linearly dependent since  
 $p_3 + p_2 - 4p_1 = 0$ .

**Solution:** On the whiteboard!

## Remarks

The main difference between linear dependence in  $\mathbb{R}^n$  and in a general vector space is that when the vectors are not  $n$ -tuples, the homogeneous equation usually cannot be written as a system of  $n$  linear equations. That is, the vectors cannot be made into the columns of a matrix  $A$  in order to study the equation  $Ax = 0$ . We must rely instead on the definition of linear dependence.

## Theorem

Let  $V$  be a vector space. Let  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  be linearly independent elements in  $V$ . Let  $\alpha_1, \dots, \alpha_k$  and  $\beta_1, \dots, \beta_k$  be numbers such that

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k = \beta_1\mathbf{v}_1 + \cdots + \beta_k\mathbf{v}_k.$$

Then we must have

$$\alpha_1 = \beta_1, \alpha_2 = \beta_2, \dots, \alpha_k = \beta_k.$$

## Proof.

On the whiteboard!



## Linear dependence and spanning set

Theorem (Spanning set theorem)

Let  $S = \{v_1, v_2, \dots, v_k\}$  be a set in a vector space  $V$ , and let  $H = \text{span}(S)$ .

- ▶ If one of the vectors in  $S$ , say  $v_k$ , is a linear combination of the remaining vectors in  $S$ , then the set formed from  $S$  by removing  $v_k$  still spans  $H$ .
- ▶ If  $H \neq \{0\}$ , then some subset of  $S$  is a basis for  $H$ .

Proof.

On the whiteboard!



# Basis and dimension I

- ▶ Definition of a basis: Let  $V$  be a vector space and  $S = \{v_1, v_2, \dots, v_n\} \subseteq V$ . If
  - (i)  $S$  spans  $V$  (i.e.,  $\text{span}(S) = V$ )
  - (ii)  $S$  is linearly independent,then  $S$  is called a **basis** of  $V$ .
- ▶ Examples:
  - i)  $\emptyset$  is a basis for  $\{\vec{0}\}$
  - ii) the standard basis for  $R^3$ :  $\{e_1, e_2, e_3\}$ , where

$$e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1).$$

- iii) Another basis for  $\mathbb{R}^3$ :  
 $v_1 = (1, 1, 1), v_2 = (1, 1, 0), v_3 = (1, 0, 0)$ .

## Basis and dimension II

iv) the standard basis for  $M_{2\times 2}$ :

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

v) The standard basis for polynomials of degree  $\leq n$ :

$$\{1, x, x^2, \dots, x^n\}.$$

vi) Another basis for polynomials of degree  $\leq n$ :

$$\{1, (1+x), (1+x)^2, \dots, (1+x)^n\}.$$

# Properties of basis

## Theorem

Let  $V$  be a vector space.

1. If  $S = \{v_1, v_2, \dots, v_n\}$  is a basis for  $V$ , then every vector in  $V$  can be written in one and only one way as a linear combination of vectors in  $S$ .
2. Let  $S$  be a basis of  $V$ . If  $|S| = n$ , then every set containing more than  $n$  vectors in  $V$  is linearly dependent.

# Maximal subset of linearly independence

A set of vectors  $S = \{v_1, v_2, \dots, v_k\}$  in  $V$  is called a **maximal** subset of linearly independent elements if  $S$  is linearly independent, and if  $S \cup \{v\}$  is linearly dependent for any  $v \notin S$ .

## Theorem

1. If  $S = \{v_1, v_2, \dots, v_k\}$  is a spanning set of  $V$  and if  $S$  is a maximal subset of linearly independent elements, then  $S$  is a basis of  $V$ .
2. Let  $V$  be a vector space and suppose that one basis has  $n$  elements, and another basis has  $m$  elements. Then  $m = n$ , i.e., all bases for a finite-dimensional vector space has the same number of vectors.

## Remarks

- When a vector space  $V$  has a basis of finite elements, then the number of vectors in a basis for  $V$  is called the **dimension** of  $V$ , denoted by  $\dim(V)$ . In other words, if  $S$  is a finite set and  $S$  is a basis of  $V$ , then

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

- A vector space  $V$  is called **finite dimensional**, if it has a basis consisting of a finite number of elements.
- If a vector space  $V$  is not finite dimensional, then it is called **infinite dimensional**.

## Remarks

1.  $\dim(\{\vec{0}\}) = 0.$
2.  $\dim(V) = n$  and  $S \subseteq V$ .
  - ▶ If  $S$  is a generating set, then  $|S| \geq n$ .
  - ▶ If  $S$  is a linear independent set, then  $|S| \leq n$ .
  - ▶ If  $S$  is a basis, then  $|S| = n$ .
3. If  $\dim(V) = n$  and if  $W$  is a subspace of  $V$ , then  $\dim(W) \leq n$ .

## Examples

1.  $\dim(\mathbb{R}^b) = n$
2.  $\dim(\mathbb{M}_{m \times n}) = mn$
3.  $\dim(P_n(x)) = n + 1$ , where  $P_n(x)$  the vector space of all polynomials with degree  $\leq n$ .
4.  $\dim(P(x)) = \infty$ , where  $P(x)$  the vector space of all polynomials.
5. Find the dimension of the subspace

$$H = \{(a - 3b + 6c, 5a + 4d, b - 2c - d, 5d) : a, b, c, d \in \mathbb{R}\}.$$

**Answer:** On the whiteboard.

## Null and column spaces

In applications of linear algebra, subspaces of  $\mathbb{R}^n$  usually arise in one two ways:

- ▶ as the set of all solutions to a homogeneous linear system
- ▶ as the set of all linear combinations of certain vectors.

These two subspaces are called null and column spaces. In the sense of linear transformation, they also called kernel and range of the linear transformation respectively.

## Null and column spaces

Given a matrix  $A$  of size  $m \times n$ .

1. **Null space** of  $A$  is the set of all solutions of the homogeneous equation  $Ax = 0$ . In set notation,

$$Nul(A) = \{x \in \mathbb{R}^n : Ax = 0\}$$

2. **Column space** of  $A$  is the set of all linear combinations of the columns of  $A$ . In the set notation,

$$Col(A) = \text{span}\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\} = \{\mathbf{b} : b = Ax \text{ for some } \mathbf{x} \in \mathbb{R}^n\},$$

where  $\mathbf{a}_i$  are columns of  $A$ .

## Theorem

Given a matrix  $A$  of size  $m \times n$ . The null space  $\text{Null}(A)$  of  $A$  is a subspace of  $\mathbb{R}^n$ . Equivalently, the set of all solutions to a homogeneous linear system with  $m$  equations,  $n$  unknown is a subspace of  $\mathbb{R}^n$ .

## Remarks

- ▶ There is no obvious relation between vectors in  $\text{Null}(A)$  and the entries in  $A$ .
- ▶ The  $\text{Null}(A)$  is defined implicitly, because it is defined by a condition that must be checked.
- ▶ No explicit list or description of the elements in  $\text{Null}(A)$  is given. However, solving the equation  $Ax = 0$  produces an explicit description of  $\text{Null}(A)$ .

## Finding a basis for $\text{Null}(A)$

1. Find the reduced echelon-form  $(A^{**}|0)$  of the augmented matrix  $(A|0)$ .
2. Write dependent variables in terms of free variables (free variables are variables corresponding to non-pivot columns of  $A^{**}$ )
3. Find general solution to  $Ax = 0$  in term of free variables.
4. Decompose the general solution vector into a linear combination of vectors where the coefficients are the free variables.
5. Vectors in the above step form a basis for  $\text{Null}(A)$ .
6. Moreover,  $\dim(\text{Null}(A))$  is equal to the number of free variables.

## Examples I

Find a spanning set and a basis for  $\text{Null}(A)$ .

1.  $A = \begin{pmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{pmatrix}$ .

**Solution:** We reduce the augmented matrix  $(A|0)$  to reduced-row echelon form:

$$\begin{aligned} A &= \left( \begin{array}{ccc|c} 1 & -3 & -2 & 0 \\ -5 & 9 & 1 & 0 \end{array} \right) \rightarrow \left( \begin{array}{ccc|c} 1 & -3 & -2 & 0 \\ 0 & -6 & -9 & 0 \end{array} \right) \rightarrow \left( \begin{array}{ccc|c} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & -6 & -9 & 0 \end{array} \right) \\ &\rightarrow \left( \begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & \frac{3}{2} & 0 \end{array} \right) \end{aligned}$$

Hence,  $x_1, x_2$  are dependent variables and  $x_3$  is free variable.  
We get

$$x = (x_1, x_2, x_3) = (-x_3, -\frac{3}{2}x_3, x_3) = x_3(-1, -\frac{3}{2}, 1)$$

and  $(-1, -\frac{3}{2}, 1)$  is a basis for  $\text{Null}(A)$ .

## Examples II

$$2. A = \begin{pmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix}.$$

**Solution:** On the whiteboard.

# Column space

## Theorem

Given a matrix  $A$  of size  $m \times n$ . Assume that  $A^*$  is the row-echelon form of  $A$ . Then, the column space of  $A$  is the space spanned by column of  $A$  corresponding to pivot columns in  $A^*$ .

**Examples:**  $A = \begin{pmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{pmatrix}$ . We have

$$A^* = \begin{pmatrix} 1 & -3 & -2 \\ 0 & -6 & -9 \end{pmatrix},$$

and

$$\{(1, -5), (-3, 9)\}$$

is a basis for  $\text{Col}(A)$ .

## Examples

Find a basis and the dimension for  $\text{Col}(A)$  where

$$A = \begin{pmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix}.$$

**Solution:** On the whiteboard.

## Row space

- ▶ Similarly, we can define the row space of a matrix, denoted by  $\text{row}(A)$ , which is a subspace generated by its rows. Note that  $\dim(\text{row}(A))$  is equal to the number of non-zero rows in the row-echelon form  $A^*$ . Furthermore, non-zero rows of  $A^*$  forms a basis for  $\text{row}(A)$ .
- ▶ Given  $A \in M_{m \times n}$ . It is straightforward that

$$\dim(\text{Col}(A)) = \dim(\text{row}(A)).$$

- ▶ We define the **rank** of a matrix  $A$ , denoted by  $r(A)$ , as the dimension of row (column) space of  $A$ .

The following is straightforward from the fact that

number of pivot columns + number of non-pivot columns = number of columns

- ▶  $r(A) + \dim(\text{null}(A)) = n$
- ▶  $r(A) + \dim(\text{null}(A^T)) = m$

## Theorem

Let  $A \in M_{n \times n}$  and  $A$  be invertible. The following are equivalent

- ▶ Columns of  $A$  form a basis of  $\mathbb{R}^n$ .
- ▶  $\text{col}(A) = \mathbb{R}^n$
- ▶  $\dim(\text{col}(A)) = n$
- ▶  $r(A) = n$
- ▶  $\text{null}(A) = \{\vec{0}\}$
- ▶  $\dim(\text{null}(A)) = 0$

## Linear algebra: Linear mappings

Lectures 7+8+9+10+11: Vector spaces

## Introduction

- ▶ Among mappings, the linear mappings are the most important.
- ▶ A good deal of mathematics is devoted to reducing questions concerning arbitrary mappings to linear mappings.
- ▶ On the other hand, it is often possible to approximate an arbitrary mapping by a linear one, whose study is much easier than the study of the original mapping. This is done in the calculus of several variables.

## Linear mappings

Let  $V, W$  be vector spaces on  $\mathbb{R}$ . A **linear mapping**

$$T : V \rightarrow W$$

is a mapping which satisfies the following two properties

- ▶  $T(u + v) = T(u) + T(v)$  for any  $u, v \in V$
- ▶  $T(cu) = cT(u)$  for any  $c \in \mathbb{R}$  and  $u \in V$ .

## Examples I

- ▶ (Projection):

$$T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$$

$$(x, y, z) \mapsto (x, y)$$

- ▶ (Inner product):

$$T : \mathbb{R}^3 \rightarrow \mathbb{R}$$

$$(x, y, z) \mapsto ax + by + cz,$$

where  $a, b, c$  are given numbers in  $\mathbb{R}$ .

- ▶ (Linear mapping given by a matrix):

$$T : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$\mathbf{v} \mapsto A\mathbf{v},$$

where  $A$  is a given matrix.

## Examples II

- ▶ (A linear transformation from  $M_{m \times n}$  to  $M_{n \times m}$ )

$$\begin{aligned}T : M_{m \times n} &\rightarrow M_{n \times m} \\ A &\mapsto A^T\end{aligned}$$

- ▶ (Differentiation):  $T$  transforms a differentiable function  $f$  to its derivative.

# Properties of linear mappings

Given a linear mapping

$$T : V \rightarrow W.$$

Then

- ▶  $T(\vec{0}) = \vec{0}$ ;
- ▶  $T(-\mathbf{v}) = -T(\mathbf{v})$
- ▶ If  $\mathbf{v} = c_1 \mathbf{v}_1 + \cdots + c_k \mathbf{v}_k$ , then

$$T(\mathbf{v}) = c_1 T(\mathbf{v}_1) + c_2 T(\mathbf{v}_2) + \cdots + c_k T(\mathbf{v}_k).$$

## Examples

Consider a map  $T$  from  $M_{m \times n}$  to  $M_{n \times m}$ :

$$\begin{aligned} T : M_{m \times n} &\rightarrow M_{n \times m} \\ A &\mapsto A^T \end{aligned}$$

Then  $T$  is a linear transformation since

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

and

$$T(cA) = (cA)^T = cA^T = cT(A).$$

## Theorem

Let  $V$  and  $W$  be vector spaces. Let  $v_1, v_2, \dots, v_n$  be a basis for  $V$ . Let  $w_1, w_2, \dots, w_n$  be arbitrary elements of  $W$ . Then there exists a unique linear mapping  $T : V \rightarrow W$  such that

$$T(v_1) = w_1, \dots, T(v_n) = w_n.$$

In other words, a linear mapping is well defined by its images of a basis.

# The kernel and image of a linear mapping

- ▶ **Kernel of a linear map  $T$ :** Let  $T : V \rightarrow W$  be a linear map. Then the set of all vectors  $v$  in  $V$  that satisfies  $T(v) = 0$  is called the **kernel** of  $T$  and is denoted by  $\ker(T)$ .

$$\ker(T) = \{v : T(v) = 0\}.$$

- ▶ **Examples:** Given

$$T(x) = Ax = \begin{bmatrix} 1 & -1 & -2 \\ -1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Then,  $\ker(T)$  coincides with solution set  $\text{null}(A)$  of  $Ax = 0$ . To calculate  $\ker(T)$ , we use Gauss-Jordan elimination as presented in the previous lecture.

## Observation

The kernel of a linear transformation  $T : V \rightarrow W$  is an essentially solution set of a homogeneous system of linear equations, and therefore is a subspace of the domain  $V$ . For this reason, sometimes  $T$  is called the **nullspace** of  $T$ .

## Some popular linear transformations

On the whiteboard.

- ▶ Composition of linear transformation
- ▶ Inverse linear transformation
- ▶ Matrix of a transformation in any bases:

## Coordinate systems

On the white board.

## Part 2: General Algebra

Topics cover:

- ▶ Propositional logic
- ▶ Relations
- ▶ Group theory
- ▶ Proof methods (if time available)

## Part 2: General Algebra

Lectures 12+13: Propositional Logics

## Statements

- ▶ A **statement** is an assertion that can be determined to be true or false.
- ▶ The **truth value** of a statement is T if it is true and F if it is false.

**Example:**

The statement " $2 + 3 = 5$ " has truth value T.

# Compound Statements

## Connectives:

- ▶ and
- ▶ or
- ▶ not
- ▶ if ... then
- ▶ ... if and only if ...

Statements that involve one or more of the connectives are **compound** statements. Otherwise they are **simple** statements.

## Examples of Compound Statements

- ▶ "If you finish your homework then you can play the game."
- ▶ "This is a question if and only if this is the correct answer."
- ▶ "I have read this and I understand the problem."

# Connectives and Symbols

Connective	Symbol	Formal name
not	$\neg$	negation
and	$\wedge$	conjunction
or	$\vee$	disjunction
if ... then	$\Rightarrow$	conditional
... if and only if...	$\Leftrightarrow$	biconditional

## Connective Or

- ▶ Note that the connective **Or** in logic is used in the inclusive sense, not the exclusive sense as in English.
- ▶ The logical statement

*"It is raining or the sun is shining"*

means

- ▶ it is raining, or
- ▶ the sun is shining, or
- ▶ it is raining and the sun is shining.

## Conditional connective

Note: There are several ways to express the conditional statement  
 $p \rightarrow q$

- ▶ If  $p$  then  $q$
- ▶  $q$  if  $p$
- ▶  $p$  is sufficient for  $q$
- ▶  $q$  is a necessary condition for  $p$
- ▶  $p$  only if  $q$

## Translating English Sentences into Propositional Sentences

Let  $p$  be the statement "The moped is still driving" and let  $q$  be the statement "The traffic light is red".

►  $p \vee q$ :

"*The moped is still driving or the traffic light is red (or both).*"

## Translating English Sentences into Propositional Sentences

Let  $p$  be the statement "The moped is still driving" and let  $q$  be the statement "The traffic light is red".

►  $q \Rightarrow p$ :

"The traffic light is red then the moped is still driving."



## Translating English Sentences into Propositional Sentences

Let  $p$  be the statement "The moped is still driving" and let  $q$  be the statement "The traffic light is red".

►  $\neg p \wedge q$ :

"The moped isn't driving anymore and the traffic light is red."

## Summary of Truth Values

- ▶ The truth value of a compound statement is determined from the truth values of its simple components under certain rules.
- ▶ These rules are summarized in a so-called **truth table**.

$p$	$\neg p$
T	F
F	T

## Truth Value

- ▶ If  $p$  and  $q$  are statements, then the truth value of the statement  $p \vee q$  is T except when both  $p$  and  $q$  have truth value F.
- ▶ The truth value of  $p \wedge q$  is F except if both  $p$  and  $q$  are true.

## Truth Table for Some Connectives

$p$	$q$	$p \vee q$	$p \wedge q$	$p \Rightarrow q$	$p \Leftrightarrow q$
T	T	T	T	T	T
T	F	T	F	F	F
F	T	T	F	T	F
F	F	F	F	T	T

## Example 1

Complete the truth table!

$p$	$q$	$p \wedge \neg q$
T	T	
T	F	
F	T	
F	F	

## Example 1

$p$	$q$	$p \wedge \neg q$	$p$	$q$	$p \wedge \neg q$
T	T	tFf	T	T	F
T	F	tTt	$\Rightarrow$		T
F	T	fFf	F	T	F
F	F	fFt	F	F	F

The lower case  $t$  and  $f$  were used to record truth values in intermediate steps.

## Example 2

Complete the truth table!

$p$	$q$	$r$	$(p \Rightarrow q)$	$\Rightarrow$	$(q \vee r)$
T	T	T	T	T	T
T	T	F	T	T	T
T	F	T	F	T	T
T	F	F	F	T	F
F	T	T	T	T	T
F	T	F	T	T	T
F	F	T	T	T	T
F	F	F	F	T	F

## Example 2 Continued

Continue to complete the truth table!

$p$	$q$	$r$	$(p \Rightarrow q) \Rightarrow (q \vee r)$
T	T	T	t
T	T	F	t
T	F	T	f
T	F	F	f
F	T	T	t
F	T	F	t
F	F	T	t
F	F	F	t

## Example 2 Continued

Continue to complete the truth table!

$p$	$q$	$r$	$(p \Rightarrow q)$	$\Rightarrow$	$(q \vee r)$
T	T	T	t	t	
T	T	F	t	t	
T	F	T	f	t	
T	F	F	f	f	
F	T	T	t	t	
F	T	F	t	t	
F	F	T	t	t	
F	F	F	t	f	

## Example 2 Continued

Continue to complete the truth table!

$p$	$q$	$r$	$(p \Rightarrow q)$	$\Rightarrow$	$(q \vee r)$
T	T	T	t	T	t
T	T	F	t	T	t
T	F	T	f	T	t
T	F	F	f	T	f
F	T	T	t	T	t
F	T	F	t	T	t
F	F	T	t	T	t
F	F	F	t	F	f

## Remark Truth Tables

- ▶ When you're constructing a truth table, you have to consider all possible assignments of True (T) and False (F) to the component statements.
- ▶ Suppose the component statements are  $P$ ,  $Q$ , and  $R$ . Each of these statements can be either true or false, so there are  $2^3 = 8$  possibilities.

## Tautology and Contradiction

- ▶ A statement that is always true is called logically true or a tautology.
- ▶ A statement that is always false is called logically false or a contradiction.

## Logical equivalence

- ▶ We cannot construct any more than 16 truth tables involving two statements.
- ▶ This is because such a truth table has 4 rows and the truth value of each row is T or F.
- ▶ However, we can certainly construct more than 16 statements involving two statements.
- ▶ What happens is that many (in fact infinitely many) statements have identical truth tables.
- ▶ We say that the statements  $r$  and  $s$  are **logically equivalent** if their truth tables are identical.

# Logical equivalence

$p$	$q$	$\neg p \vee q$
T	T	T
T	F	F
F	T	T
F	F	T

This is equivalent to the truth table of  $p \Rightarrow q$ .

# Logical equivalence

- ▶ The statements  $r$  and  $s$  are equivalent if and only if  $r \Leftrightarrow s$  is a tautology.

# Implication

- ▶ We say that  $r$  implies  $s$  if  $s$  is true whenever  $r$  is true.
- ▶ If  $r$  implies  $s$  then we write  $r \Rightarrow s$ .
- ▶ Alternatively,  $r$  implies  $s$  if and only if the statement  $r \Rightarrow s$  is a tautology.
- ▶ We say that  $s$  is logically deducible from  $r$ .
- ▶ For example, in a mathematical theorem the hypothesis implies the conclusion or the conclusion is deducible from the hypothesis.

# Order of Precedence

Operator	Precedence
$\neg$	1
$\wedge$	2
$\vee$	3
$\Rightarrow$	4
$\Leftrightarrow$	5

# Order of Precedence

- $\neg$  has higher precedence than  $\wedge$ , i.e.

$$\neg p \wedge q$$

should be read

$$(\neg p) \wedge q$$

and not

$$\neg(p \wedge q)$$

# Order of Precedence

- $\wedge$  has higher precedence than  $\vee$ , i.e.

$$p \wedge q \vee r$$

should be read

$$(p \wedge q) \vee r$$

and not

$$p \wedge (q \vee r)$$

# Order of Precedence

- ▶  $\vee$  has higher precedence than  $\Rightarrow$ , i.e.

$$p \vee q \Rightarrow r$$

should be read

$$(p \vee q) \Rightarrow r$$

and not

$$p \vee (q \Rightarrow r)$$

# Order of Precedence

- $\Rightarrow$  has higher precedence than  $\Leftrightarrow$ , i.e.

$$p \Rightarrow q \Leftrightarrow r$$

should be read

$$(p \Rightarrow q) \Leftrightarrow r$$

and not

$$p \Rightarrow (q \Leftrightarrow r)$$

## Commutative Laws

- ▶  $p \wedge q \Leftrightarrow q \wedge p$
- ▶  $p \vee q \Leftrightarrow q \vee p$

## Associative Laws

- $p \wedge (q \wedge r) \Leftrightarrow (p \wedge q) \wedge r$
- $p \vee (q \vee r) \Leftrightarrow (p \vee q) \vee r$

## Distributive Laws

- $p \wedge (q \vee r) \Leftrightarrow p \wedge q \vee p \wedge r$
- $p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$

## Idempotent (or Tautology)

- ▶  $p \wedge p \Leftrightarrow p$
- ▶  $p \vee p \Leftrightarrow p$

## Absorption

- $p \wedge (p \vee q) \Leftrightarrow p$
- $p \vee (p \wedge q) \Leftrightarrow p$

## Complementation

- $p \wedge \neg p \Leftrightarrow F$
- $p \vee \neg p \Leftrightarrow T$

## Law of Involution (Double Complementation)

$$\blacktriangleright \neg\neg p \Leftrightarrow p$$

## Laws of de Morgan

- ▶  $\neg p \wedge \neg q \Leftrightarrow \neg(p \vee q)$
- ▶  $\neg(p \wedge q) \Leftrightarrow \neg p \vee \neg q$

## Identity Elements

Disjunction:

- ▶  $p \vee T \Leftrightarrow T$
- ▶  $p \vee F \Leftrightarrow p$

Conjunction:

- ▶  $p \wedge T \Leftrightarrow p$
- ▶  $p \wedge F \Leftrightarrow F$

# Logic gates

Some popular logic gates

- ▶ Negation gate
- ▶ AND gate
- ▶ OR gate
- ▶ NAND gate
- ▶ NOR gate
- ▶ XOR gate

Details are on the white board.

## Rules of Inference

- ▶ An argument is a sequence of statements that end with a **conclusion**. An argument is **valid** if the conclusion follows from the truth of the preceding statements (**premises** or **hypotheses**).
- ▶ In propositional logic, an argument is valid if it is based on a tautology.
- ▶ Arguments that are not based on tautology are called **fallacies**.

Name	Rule of Inference	Tautology
Addition	$\begin{array}{c} p \\ \therefore \underline{p \vee q} \end{array}$	$p \rightarrow (p \vee q)$
Simplification	$\begin{array}{c} p \wedge q \\ \therefore p \end{array}$	$(p \wedge q) \rightarrow p$
Modus ponens	$\begin{array}{c} p \\ p \rightarrow q \\ \therefore q \end{array}$	$p \wedge (p \rightarrow q) \rightarrow q$
Modus tollens	$\begin{array}{c} \neg q \\ p \rightarrow q \\ \therefore \neg p \end{array}$	$(\neg q) \wedge (p \rightarrow q) \rightarrow \neg p$
Hypothetical syllogism	$\begin{array}{c} p \rightarrow q \\ q \rightarrow r \\ \therefore p \rightarrow r \end{array}$	$(p \rightarrow q) \wedge (q \rightarrow r) \rightarrow (p \rightarrow r)$
Disjunctive syllogism	$\begin{array}{c} \neg p \\ p \vee q \\ \therefore q \end{array}$	$(p \vee q) \wedge (\neg p) \rightarrow q$

Given the hypotheses:

- ▶ "It is not sunny and is cold"
- ▶ "We go swimming only if it is sunny"
- ▶ "If we do not go swimming then we will play soccer"
- ▶ "If we play soccer then we will go home by sunset"

Show that these hypotheses lead to the conclusion: "We will go home by sunset".

Given the hypotheses:

- ▶ "If you send me an email, I will finish writing the program"
- ▶ "If you do not send email then I will go to bed early"
- ▶ "If I go to bed early then I will go jogging tomorrow morning"

Show that these hypotheses lead to the conclusion: "If I do not finish writing the program then I will go jogging tomorrow morning".

## Some fallacies

- ▶ Fallacy of affirming the conclusion:  $[(p \rightarrow q) \wedge q] \rightarrow p$
- ▶ Fallacy of denying the hypothesis:  $[(p \rightarrow q) \wedge \neg p] \rightarrow \neg q$

# Variable

- ▶ The sentence " $x + 2 = 5$ " cannot be assigned a truth value and so it is technically not a statement.
- ▶ However, if we knew the value of  $x$  we could determine its truth value.
- ▶ We say that it is a statement involving a **variable**, namely  $x$ .
- ▶ If we allow  $x$  to have value 3 then we get a true statement whereas, if we let  $x$  have value 1 we get a false statement.

## Quantifiers

- ▶ Another way to be able to assign truth values to statements involving variables is through the use of **quantifiers**.
- ▶ The **universal** quantifier,  $\forall$ , in English means "For all possible values of the variable".
- ▶ The **existential** quantifier,  $\exists$ , in English means "There is a value of the variable".

## Quantifiers

- ▶  $\forall x, x + 2 = 5$
- ▶ This states that for all possible values of the variable  $x$  that  $x + 2 = 5$ .

## Quantifiers

- ▶  $\exists x, x + 2 = 5$
- ▶ This states that there is a value of  $x$  for which  $x + 2 = 5$ .

# Quantifiers

- ▶ The problem we are left with is what do we mean by the possible values of the variable?
- ▶ We assume that the variable in our statement may be replaced by a quantity from a set of known objects.
- ▶ This set of objects is called the **universal set of discourse**, or more briefly, **the universal set**.
- ▶ Usually, this universal set will be understood from the context.
- ▶ If necessary, we shall define the universal set for each variable in the statement, e.g.  $\forall x \in U, x + 2 = 5$ .
- ▶ This translates into English as " $x$  an element of the set  $U$ ".

## Truth Set

- ▶ Alternatively, we may think about a statement involving variables as a statement about the universal set  $U$ .
- ▶ The truth set of the statement consists of those elements of  $U$  for which the statement is true.
- ▶ For example, if the universal set is the set of integers then the truth set of the statement

$$x^2 = 4$$

is -2 and 2.

## Truth Set

- ▶ A statement about  $U$  is a tautology if and only if its truth set is  $U$  and a contradiction if its truth set is the empty set.

## Quantifiers and the Truth Set

- ▶ Thus a statement such as  $\forall x, f(x)$  is true if its truth set is  $U$ .
- ▶ If its truth set is a proper subset of  $U$  then it is a false statement.
- ▶ A statement of the form  $\exists x, f(x)$  is true if its truth set is non empty and false if its truth set is the empty set.

# Negation

- ▶ Consider the statement  $\forall x, f(x)$ .
- ▶ If this statement is true then its truth set is  $U$ .
- ▶ Its negation must be false and have truth set the empty set.
- ▶ Alternatively, if the statement is false (i.e. its truth set is a proper subset of  $U$ ) then its negation has to be true.
- ▶ We claim that the negation is the statement  $\exists x, \neg f(x)$ .
- ▶ Clearly this is the negation if the original statement is true.

## Negation

- ▶ If the original is false then there must have been some  $x$  for which  $f(x)$  was false.
- ▶ We conclude that the statement

$$\neg(\forall x, f(x))$$

is equivalent to

$$\exists x, \neg f(x).$$

- ▶ Similarly,

$$\neg(\exists x, f(x))$$

is equivalent to

$$\forall x, \neg f(x).$$

# Negation

Statement	Negation
$\forall x, \forall y, f(x, y)$	$\exists x, \exists y, \neg f(x, y)$
$\forall x, \exists y, f(x, y)$	$\exists x, \forall y, \neg f(x, y)$
$\exists x, \forall y, f(x, y)$	$\forall x, \exists y, \neg f(x, y)$
$\exists x, \exists y, f(x, y)$	$\forall x, \forall y, \neg f(x, y)$

## Part 2: General Algebra

Lecture 14: Set theory review

# Sets

- ▶ A set  $L$  is a collection of objects, called *elements* of the set. A set can be represented by listing its elements between braces:

$$\text{Ex.: } L = \{a, b, c, d\}$$

- ▶ Notation:
  - ▶ An element belongs to a set:  $\in$
  - ▶ An element does not belong to a set:  $\notin$

$$\text{Ex.: } a \in L \text{ and } z \notin L$$

# Sets

- ▶ Recurrence and order is not important:

*Bsp.:*

$$\{ \text{red} , \text{blue} , \text{red} \} = \{ \text{red}, \text{blue} \} \text{ und}$$

$$\{ 3, 1, 9 \} = \{ 9, 3, 1 \} = \{ 1, 3, 9 \}$$

- ▶ Two sets  $A$  and  $B$  are equal if and only if they have the same elements.

*Ex.:  $A = B$*

# Sets

- ▶ The elements must not have anything in common or be especially similar:

*Ex.:  $L = \{a, 7, \{c, \text{red}\}\}$*

- ▶ A set with no elements is called *empty set*:

$$\emptyset = \{\}$$

- ▶ A set  $A$  is a subset of  $B$ , if all elements of  $A$  are in  $B$ :

$$A \subseteq B$$

# Sets

- ▶ We call  $A$  a proper subset of  $B$  if  $A \subseteq B$  but  $A \neq B$ , i.e.  $A \subset B$ , i.e. there is some element in  $B$  which is not in  $A$ .
- ▶  $A = B$  if and only if  $A \subseteq B$  and  $B \subseteq A$
- ▶  $\emptyset \subseteq$  all sets
- ▶ If the set is finite, its number of elements is represented by  $|A|$

## Characterization of Elements

- ▶ Infinite sets are written with dot notation:

*Ex.: The set of natural numbers,  $\mathbb{N} = \{0, 1, 2, \dots\}$*

- ▶ To specify a certain property for the elements of the set:

$B = \{x \mid x \text{ has property } P_1 \text{ and } P_2 \text{ and } \dots\}$

*Ex.: The set of odd natural numbers:*

$B = \{x \mid x \in \mathbb{N} \text{ and } x \text{ not divisible by } 2\}$

# Complement

► Complement:

Let  $A$  be a subset of the universal set  $\Omega$ , i.e.  $A \subseteq \Omega$ .

The complement  $\bar{A}$  of  $A$  in  $\Omega$  is the set of elements that do not belong to  $A$ :

$$\bar{A} = \{x \mid x \in \Omega \text{ and } x \notin A\}$$

## Union

- ▶ Union:

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

i.e. the set of elements that belong to either of  $A$  and  $B$ .

Ex.:  $\{1, 2, 3\} \cup \{3, 4, 5\} = \{1, 2, 3, 4, 5\}$

## Intersection

- ▶ Intersection:

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

i.e. the result is the set containing the common elements of two sets:

Ex.:  $\{1, 2, 3\} \cap \{3, 4, 5\} = \{3\}$

## Difference (or Relative Complement)

- ▶ Difference:

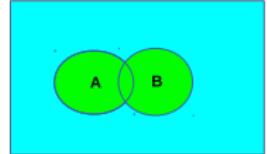
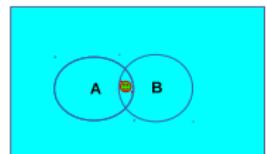
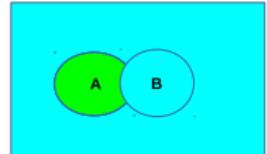
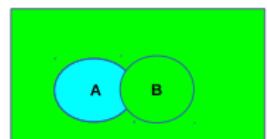
$$A - B = A \setminus B = \{x \mid x \in A \text{ and } x \notin B\},$$

i.e. the result is the set of elements that belong to a set but not to another:

$$\text{Ex.: } \{1, 2, 3\} - \{3, 4, 5\} = \{1, 2\}$$

- ▶ The sets  $A$  and  $B$  are disjoint, if  $A \cap B = \emptyset$ , i.e. if they have no common elements.

## Example: Sets as Venn Diagrams

$A \cup B$	
$A \cap B$	
$A \setminus B$	
	

# Cartesian Product

- ▶ The Cartesian product of two sets  $A$  and  $B$  is the set, which elements are ordered pairs  $(a, b)$ , with  $a \in A$  and  $b \in B$ :

$$A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}$$

- ▶ *Two examples:*

If  $\mathbb{R}$  are the real numbers, then the Cartesian product  $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$  is the real plane.

The set  $\mathbb{R} \times \mathbb{R} \times \mathbb{R} = \mathbb{R}^3$  is the real space.

## Example: Cartesian Product

Let  $A = \{1, 2, 3\}$  and  $B = \{a, b\}$ . Then is

$$A \times B = \{(1, a), (1, b), (2, a), (2, b), (3, a), (3, b)\}$$

Thus the Cartesian product has  $3 \cdot 2 = 6$  elements.

# Cartesian Product

- ▶ Notice: If  $A \neq B$ , then we have:  $A \times B \neq B \times A$
- ▶ For finite sets  $A$  and  $B$  we have:

$$|A \times B| = |A| \cdot |B|,$$

since we can choose the first coordinate in the pair in  $|A|$  ways and the second in  $|B|$  ways.

The multiplication principle gives that the number of elements of  $A \times B$  equals  $|A| \cdot |B|$ .

## Computer representation of sets I

Let  $U$  be a universal set. Fix an arbitrary order of elements of  $U$  for instance  $a_1, a_2, \dots, a_n$ .

If  $A$  is a subset of  $U$ , represent  $A$  with a bit string of length  $n$ , where the  $i$ th bit is 1 if  $a_i$  is in  $A$ , and is 0 if  $a_i$  is not in  $A$ .

**Example.** Let  $U = \{1, 2, 3, 4, 5, 6, 7, 8\}$ . Then the subset

- ▶  $A = \{1, 3, 4, 6\}$  is represented as the bit string 10110100
- ▶  $B = \{4, 1, 3, 5\}$  is represented as the bit string 10111000
- ▶  $A \cap B$  is represented as the bit string (please check!):

$$10110100 \wedge 10111000 = 10110000.$$

Hence,  $A \cap B = \{1, 3, 4\}$ .

- ▶  $A \cup B$  is represented as the bit string (please check!):

$$10110100 \vee 10111000 = 10111100.$$

Hence,  $A \cup B = \{1, 3, 4, 5, 6\}$ .

## Computer representation of sets II

### Exercises:

1. Propose an algorithm to find the union, intersection, symmetric difference of the two given sets. What is the running time (complexity) of your algorithm?
2. Using the binary representation of subsets to prove that if  $|A| = n$  then the power set  $\mathcal{P}(A)$  of  $A$  is of cardinality  $2^n$ .

## Part 2: General Algebra

Lecture 15+16+17: Relations

# Rosen's textbook: Introduction

## **Topics covered:**

- ▶ Relations and Their Properties
- ▶  $n$ -ary Relations and Their Applications
- ▶ Relation representations
- ▶ Closures of Relations
- ▶ Equivalence Relations

# What Is a Relation?

- ▶ Let  $A$  and  $B$  be sets. A (binary) relation from  $A$  to  $B$  is a subset of  $A \times B$ .
- ▶ A subset  $A^2 = A \times A$  is called a relation in  $A$ .

## Example: Relation

For example is this a relation:

>

i.e.

$$\{(i,j) | i, j \in \mathcal{N} \text{ and } i < j\}$$

# Properties of Relations

Let  $R$  be a relation in set  $A$ .

- Reflexive:

$R \subseteq A \times A$  is reflexive, if  $(a, a) \in R$  for all  $a \in A$

Ex.:  $\leq$

- Symmetric:

$R \subseteq A \times A$  is symmetric, if and only if  $(b, a) \in R$  for  $(a, b) \in R$

Bsp.: Brothers and sisters

# Properties of Relations

Let  $R$  be a relation in  $A$ .

- ▶ Antisymmetric:

$R$  is antisymmetric, if and only if  $(b, a) \notin R$ , if  $(a, b) \in R$   
(and  $a$  and  $b$  differ)

*Ex.: Mother and son*

- ▶ Transitive:

If  $(a, b) \in R$  and  $(b, c) \in R$ , then we also have  $(a, c) \in R$

*Ex.: Forefathers (Child, Father, Grandfather)*

# What is an Equivalence Relation?

- ▶ An equivalence relation is a relation, which is reflexive, symmetric and transitive.
- ▶ Such a relation partitions the set  $A$  in disjoint subsets, called equivalence classes.
- ▶ *Ex.:*

$A = \text{the set of different kinds of light bulbs.}$

One equivalence class:      40 W-light bulbs,  
another equivalence class:      60 W-light bulbs.

## Example: Equivalence Relation

- ▶ The relation *to be of the same age as somebody else* is an equivalence relation. What are the different equivalence classes?

## Example: Equivalence Classes



# Operations on Zero-One Matrices I

- ▶ A matrix  $A = (a_{ij})_{m \times n}$  is called a **zero-one matrix** if  $a_{ij} \in \{0, 1\}$  for all  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, n$ .
- ▶ Zero-one matrices are used to represent discrete structure like graphs, relation in computer networks, network flow;
- ▶ Zero-one matrices are used to represent black-white pixels in image processing.

Let  $A$  and  $B$  be two 0-1 matrices of the same sizes:

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad B = [b_{ij}] = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}.$$

## Operations on Zero-One Matrices II

- ▶ The meet of  $A$  and  $B$  is  $A \wedge B = [a_{ij} \wedge b_{ij}]$
- ▶ The join of  $A$  and  $B$  is  $A \vee B = [a_{ij} \vee b_{ij}]$
- ▶ The symmetric difference of  $A$  and  $B$  is  $A \oplus B = [a_{ij} \oplus b_{ij}]$
- ▶ The complement of  $A$  is  $\bar{A} = [\bar{a}_{ij}]$

Example. Let  $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$  and  $B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$ . Then

- ▶  $A \wedge B = \begin{bmatrix} 1 \wedge 0 & 0 \wedge 1 & 1 \wedge 0 \\ 0 \wedge 1 & 1 \wedge 1 & 0 \wedge 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$
- ▶  $A \vee B = \begin{bmatrix} 1 \vee 0 & 0 \vee 1 & 1 \vee 0 \\ 0 \vee 1 & 1 \vee 1 & 0 \vee 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$
- ▶  $\bar{A} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

## Boolean product $A \odot B$ |

Let  $A$  be an  $m \times n$  zero-one matrix,  $B$  be an  $n \times p$  zero-one matrix. The Boolean product of  $A$  and  $B$ , denoted by  $A \odot B$ , is an  $m \times p$  zero-one matrix with the entry  $c_{ij}$  is defined by

$$c_{ij} = (a_{i1} \wedge b_{1j}) \vee (a_{i2} \wedge b_{2j}) \cdots \vee (a_{in} \wedge b_{nj}).$$

**Example.**

$$\begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \odot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = [(0 \wedge 1) \vee (1 \wedge 0) \vee (0 \wedge 1)] = [0]$$

and

$$\begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \odot \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = [(0 \wedge 0) \vee (1 \wedge 1) \vee (0 \wedge 1)] = [1]$$

## Boolean product $A \odot B$ ||

**Example.**

$$\begin{aligned} & \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \odot \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} (0 \wedge 1) \vee (1 \wedge 0) \vee (0 \wedge 1) & (0 \wedge 1) \vee (1 \wedge 1) \vee (0 \wedge 0) \\ (1 \wedge 1) \vee (1 \wedge 0) \vee (0 \wedge 1) & (1 \wedge 1) \vee (1 \wedge 1) \vee (0 \wedge 0) \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \end{aligned}$$

# Boolean powers I

## Definition

Let  $A$  be a square zero-one matrix of size  $n \times n$ . The  $r$ th Boolean power of  $A$ , denoted by  $A^{[r]}$ , is the matrix

$$A^{[r]} = A \odot A \odot \cdots \odot A.$$

By convention,  $A^{[0]} = I_n$ .

Example: Let  $A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}$ . Then (please verify)

$$\blacktriangleright A^{[2]} = A \odot A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

$$\blacktriangleright A^{[3]} = A^{[2]} \odot A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

## Boolean powers II

$$\blacktriangleright A^{[4]} = A^{[3]} \odot A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$\blacktriangleright A^{[5]} = A^{[4]} \odot A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

# Combining Relations

Let  $R$  and  $S$  be relations from  $A$  to  $B$ .

- ▶  $R \cup S$  = Union relation of  $R$  and  $S$
- ▶  $R \cap S$  = Intersection relation of  $R$  and  $S$
- ▶  $R \oplus S$  = Exclusive or relation of  $R$  and  $S$
- ▶  $R - S$  = Difference relation of  $R$  and  $S$
- ▶  $\bar{R}$  = Complementary relation of  $R$  (in  $A \times B$ )
- ▶  $R^{-1}$  = Inverse relation of  $R$ , which consists of all pairs  $(b, a)$  where  $(a, b) \in R$
- ▶ Composite of two relations: Let  $R$  and  $S$  be relations

$$A \xrightarrow{R} B \xrightarrow{S} C$$

The composition  $S \circ R$  is the set of all pairs  $(a, c)$  such that there exists  $b \in B$  with  $(a, b) \in R$  and  $(b, c) \in S$ .

# Properties of Relations

Let  $A$  be a set. A relation from  $A$  to  $A$  is called a **relation on  $A$** .

Let  $R$  be a relation on  $A$ .

- ▶  $R$  is called **reflexive** if  $R$  contains all pairs  $(a, a)$  where  $a \in A$ .  
In other words,  $R$  is reflexive on  $A$  if every element in  $A$  is related to itself by  $R$ .
- ▶  $R$  is called **symmetric** if whenever  $(a, b) \in R$  then  $(b, a) \in R$ .
- ▶  $R$  is called **antisymmetric** if there do not exist  $a \neq b$  such that both  $(a, b)$  and  $(b, a)$  are in  $R$ .
- ▶  $R$  is called **transitive** if whenever  $(a, b) \in R$  and  $(b, c) \in R$  then  $(a, c) \in R$ .

**Example.** Check if the following relations are reflexive, symmetric, antisymmetric and transitive

- (1)  $R$  = set of pairs of students in a class that have the same birthday
- (2)  $R$  = set of pairs of two integers  $(a, b)$  with  $a$  divisible by  $b$
- (3)  $R$  = set of pairs of real numbers  $(x, y)$  with  $x + y = 0$
- (4)  $R$  = set of pairs of negative numbers  $(a, b)$  with  $a > 2b$

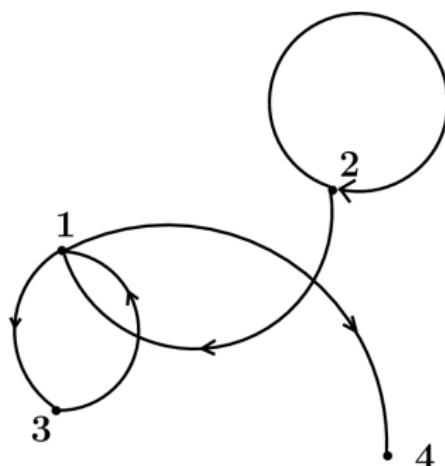
**Note.** The relation  $R$  is transitive if and only if  $R^2 \subseteq R$ . (Explain!)

## Relation representations

**Relations via directed graphs.**

Let  $R = \{(1, 3), (2, 2), (1, 4), (2, 1), (3, 1)\}$  on  $A = \{1, 2, 3, 4\}$ .

We can use digraphs to represent this relation.



**Question.** What are the properties of digraphs representing relations which are

- ▶ reflexive
- ▶ symmetric
- ▶ antisymmetric
- ▶ transitive

## Relations via 0 – 1 matrices

Given a relation  $R$  from a set of  $m$  elements  $\{a_1, a_2, \dots, a_m\}$  to a set of  $n$  elements  $\{b_1, b_2, \dots, b_n\}$ . To represent  $R$  we use a matrix of size  $m \times n$ , denoted by  $M_R = [a_{ij}]$ , whose entries are defined as follows:

$$a_{ij} = \begin{cases} 1 & \text{if } (a_i, b_j) \in R, \\ 0 & \text{if } (a_i, b_j) \notin R \end{cases}$$

**Example:** Let

$R = \{(apple, sour), (orange, sweet), (apple, sweet), (kiwi, sour)\}$   
be a relation from  $\{apple, orange, kiwi\}$  to  $\{sour, sweet, bitter\}$ .  
Then, the boolean matrix  $M_R$  for  $R$  is

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

# Operations on relations and operations on boolean matrices

## Theorem

Let  $R$  and  $S$  be relations on  $A$  with their representing matrices  $M_R$  and  $M_S$ . Then:

- ▶  $M_{R \cup S} = M_R \vee M_S$
- ▶  $M_{R \cap S} = M_R \wedge M_S$
- ▶  $M_{R \oplus S} = M_R \oplus M_S$
- ▶  $M_{\overline{R}} = \overline{M_R}$
- ▶  $M_{R^{-1}} = (M_R)^T$
- ▶  $M_{S \circ R} = M_R \odot M_S$
- ▶  $M_{R^n} = M_R \odot M_R \cdots \odot M_R =: M_R^{[n]}$

## Proof.

On the whiteboard.



**Question 1.** What are the properties of the representing matrix of a relation which is:

- ▶ reflexive
- ▶ symmetric
- ▶ antisymmetric
- ▶ transitive

**Question 2.** Which properties: reflexive, symmetric, antisymmetric, or transitive that are possessed by the relation represented by the following matrix:

$$(a) \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$(b) \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

**Question 3.** Let  $A$  be a set of cardinality  $n$ . How many relations on  $A$  are there that are:

- ▶ reflexive?
- ▶ symmetric?
- ▶ reflexive and symmetric?

## Closures of Relations

Let  $R$  be a relation on  $A$ . The relation obtained from  $R$  by adding a minimal number of new pairs so that the new relation is reflexive (symmetric, antisymmetric, transitive) is called the **reflexive closure** (symmetric, antisymmetric, transitive closure) of  $R$ .

**Example 1.** Find the reflexive closure of:

- ▶  $R = \{(1, 2), (2, 3), (3, 3)\}$  on  $A = \{1, 2, 3\}$ .
- ▶  $R = \{(a, b) | a \neq b\}$  on the set of real numbers.
- ▶  $R = \{(a, b) | a \geq b\}$  on the set of real numbers.

**Example 2.** Find the symmetric closure of:

- ▶  $R = \{(1, 2), (2, 3), (3, 3)\}$  on  $A = \{1, 2, 3\}$ .
- ▶  $R = \{(a, b) | a > b\}$  on the set of real numbers.
- ▶  $R = \{(a, b) | a \neq b\}$  on the set of real numbers.

**Example 3.** Find the antisymmetric closure of

$R = \{(1, 2), (1, 3), (2, 2), (2, 1)\}$  on  $\{1, 2, 3\}$ .

## Transitive Closure

Let

$$R = \{(1, 3), (1, 4), (2, 1), (3, 2)\}$$

on the set  $A = \{1, 2, 3, 4\}$ .

The relation  $R$  is not transitive.

To get a transitive relation from  $R$ , we first need to add the pairs of  $R^2$ .

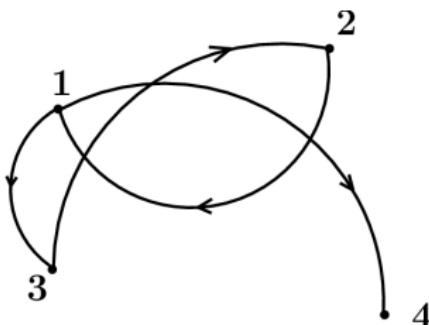
These new pairs together with those of  $R$  may form new ones, meaning we might need to add pairs of  $R^3, \dots$

Denote  $R^*$  the transitive closure of  $R$ . Then  $R^*$  is the infinite union

$$R^* = R \cup R^2 \cup R^3 \cup \dots$$

## Connectivity Relations

Let  $R = \{(1, 3), (1, 4), (2, 1), (3, 2)\}$  on  $A = \{1, 2, 3, 4\}$



- ▶  $R^n$  = set of pairs of vertices  $(u, v)$  for which there is a path of length  $n$  from  $u$  to  $v$ .
- ▶  $R \cup R^2 \cup \dots \cup R^n$  = set of pairs of vertices  $(u, v)$  for which there is a path of length at most  $n$  from  $u$  to  $v$ .
- ▶  $R^* = R \cup R^2 \cup R^3 \cup \dots$  = set of pairs of vertices  $(u, v)$  for which there is a path from  $u$  to  $v$ .

# An Algorithm for Computing Transitive Closures

Let  $R$  be a relation on a set of  $n$  elements. It is showed that to compute  $R^*$  we only need to compute a finite union of relations.

## Theorem

Let  $R$  be a relation on a set of  $n$  elements. Then

$$R^* = R \cup R^2 \cup R^3 \cup \dots \cup R^n.$$

Equivalently, the matrix of  $R^*$  is determined from the following equation

$$M_{R^*} = M_R \vee M_R^{[2]} \vee M_R^{[3]} \vee \dots \vee M_R^{[n]}$$

**Example.** Find the transitive closure of

$R = \{(1, 3), (1, 4), (2, 1), (3, 2)\}$  on  $A = \{1, 2, 3, 4\}$ .

$$M_R = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad M_R^{[2]} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$M_R^{[3]} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad M_R^{[4]} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\text{Then } M_{R^*} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

## Warshall's Algorithm for Computing Transitive Closures

Let  $R$  be a relation on a set of  $n$  elements. To compute the transitive closure of  $R$ , Warshall's algorithm constructs a sequence of matrices  $M_0, M_1, \dots$  recursively:

- ▶  $M_0 := M_R$
- ▶  $M_k := M_{k-1} \vee$  (Boolean product of  $k$ th column and  $k$ th row of  $M_{k-1}$ )

The algorithm terminates when  $k = n$ , and the matrix  $M_n$  is the matrix of the transitive closure of  $R$ .

**Example.** Use Warshall's algorithm to compute the transitive closure of  $R = \{(1, 3), (1, 4), (2, 1), (3, 2)\}$  on  $A = \{1, 2, 3, 4\}$ .

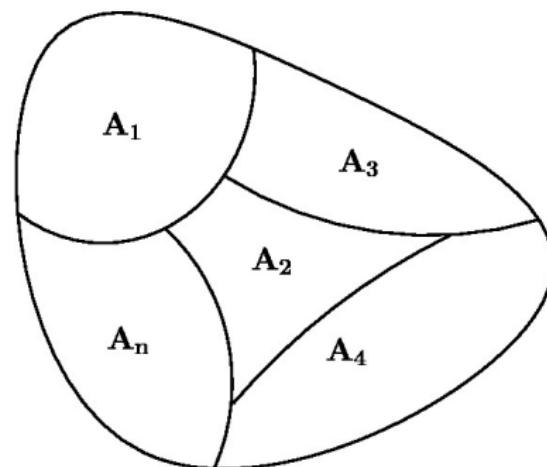
**Exercise.** Read and understand the proof of correctness of Warshall's algorithm in textbook.

# Equivalence relations

## Partitions.

Let  $A$  be a set. The subsets  $A_1, A_2, \dots, A_n$  are called a partition of  $A$  if:

- (i)  $A_1 \cup A_2 \cup \dots \cup A_n = A$
- (ii)  $A_i \cap A_j = \emptyset$  for all  $i \neq j$ .



**Example.** Which of the following subsets form a partition of the set of all real numbers?

- (a) The subset of positive integers, the subset of negative integers
- (b) The subset of non-positive integers, the subset of non-negative integers
- (c) The subset of rational numbers, the subset of irrational numbers
- (d) The closed intervals  $[n, n + 1]$  where  $n$  is an integer
- (e) The intervals  $(n, n + 1]$  where  $n$  is an integer.

# Equivalence Relations

The relation  $R$  on  $A$  is called an equivalence relation if it is reflexive, symmetric and transitive.

**Example.** The following relations are equivalence relations:

- (a) The congruence modular  $m$  relation on the set of integers
- (b) The born-in-the-same-city relation in a class
- (c) The relation whose representing matrix is  $M_R = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
- (d)  $R = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 3), (4, 4)\}$  on  $A = \{1, 2, 3, 4\}$

## Theorem

Let  $R$  be an equivalence relation on  $A$ . Then

- ▶ The set  $A$  is partitioned into disjoint subsets, elements in each subset are mutually related to each other (and to themselves), and any two elements of two different subsets are not related by  $R$ .
- ▶ These subsets are called **equivalence classes** of  $R$ .
- ▶ Therefore, an equivalence relation on  $A$  forms a partition of  $A$  consisting of equivalence classes of  $R$ .

**Example.** Find all equivalence classes of each equivalence relation

(a) The congruence modular  $m$  relation of the set of integers

(b) The born-in-the-same-city relation in a class

(c) The relation whose representing matrix is  $M_R = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(d)  $R = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 3), (4, 4)\}$  on  
 $A = \{1, 2, 3, 4\}$

Let  $A_1, A_2, \dots, A_n$  be a partition of  $A$ . There exists an equivalence relation  $R$  on  $A$  whose equivalence classes are the subsets of this partition.

**Example.** Define an equivalence relation whose equivalence classes are the subsets of the following partitions:

- ▶ Partition  $\{1, 3, 5\}, \{2\}, \{4\}$  of  $\{1, 2, 3, 4, 5\}$
- ▶ Partition of the set of integers consisting of the subset of even numbers and the subset of odd numbers

**Question.** How many equivalence relations are there on the set  $\{1, 2, 3, 4\}$ ?

# *n*-ary Relations and Their Applications

An *n*-ary relation on the sets  $A_1, A_2, \dots, A_n$  is a subset of  $A_1 \times A_2 \times \dots \times A_n$

**Example.** Given the sets

$$A_1 = \{\text{NamNT}, \text{TrungTT}, \text{HuongNTQ}, \text{ThaoNP}, \text{HienPQ}\},$$

$$A_2 = \{00198, 00011, 00345, 00786, 00321, 00546\},$$

$$A_3 = \{\text{MAD111}, \text{MAA101}, \text{MAS291}, \text{PFC111}, \text{DSA1}\}$$

$$A_4 = \{3.5, 4.0, 5.7, 8.0, 9.5, 6.4\}$$

A relation  $R$  on these sets can be expressed as a database

Name	Code	Subject	Grade
HuongNTQ	00345	MAS291	8.0
TrungTT	00786	DSA1	3.5
HienPQ	00321	DSA1	8.0

Name	Code	Subject	Grade
HuongNTQ	00345	MAS291	8.0
TrungTT	00786	DSA1	3.5
HienPQ	00321	DSA1	8.0

We say  $R$  is a relation on 4 **domains** (Name, Code, Subject and Grade).

An element of this relation is a **record** consisting of 4 **fields**. For example

(TrungTT, 00786, DSA1, 3.5)

## Primary key - Composite key

- ▶ A domain of an  $n$ -ary relation is a **primary key** if the value from this domain of the  $n$ -tuple determines this  $n$ -tuple.
- ▶ A set of domains of an  $n$ -ary relation is a **composite key** if they determine uniquely  $n$ -tuples.

**Example.** In the 4-ary relation,

Name	Code	Subject	Grade
HuongNTQ	00345	MAS291	8.0
TrungTT	00786	DSA1	3.5
HienPQ	00321	DSA1	8.0

- ▶ Each domain **Name** or **Code** can be primary key.
- ▶ Each domain **Subject** or **Grade** is not primary key, but together they form a composite key.

# Operations on $n$ -ary Relations

**Selection operator  $S_C$ .** Given relation

Name	Code	Subject	Grade
HuongNTQ	00345	MAS291	8.0
TrungTT	00786	DSA1	3.5
HienPQ	00321	DSA1	8.0

- If  $C$  is the condition ( $\text{Subject} = \text{"DSA1"}$ ) then the selection  $S_C$  produces a 4-ary relation with 2 records

Name	Code	Subject	Grade
TrungTT	00786	DSA1	3.5
HienPQ	00321	DSA1	8.0

- If  $C$  is the condition ( $\text{Subject} = \text{"DSA1"} \wedge (\text{Grade} > 5.0)$ ) then the selection  $S_C$  produces a relation with how many records?

**Projection operator  $P$ .** Given the relation

Name	Code	Subject	Grade
HuongNTQ	00345	MAS291	8.0
TrungTT	00786	DSA1	3.5
HienPQ	00321	DSA1	8.0

- ▶ The projection  $P_{2,3,4}$  produces a 3-ary relation with 3 records

Code	Subject	Grade
00345	MAS291	8.0
00786	DSA1	3.5
00321	DSA1	8.0

- ▶ The projection  $P_{2,4}$  on the above relation will produce a relation of how many domains and how many records?

**Join operator  $J_2$** . Given two relations

Name	Code	Subject
HuongNTQ	00345	MAS291
NamNT	00011	MAD121
TrungTT	00786	DSA1
HienPQ	00321	DSA1

Code	Subject	Grade
00345	MAS291	8.0
00786	DSA1	3.5
00546	MAD121	5.7
00321	DSA1	8.0

The join operator  $J_2$  used on these relations will produce the relation

Name	Code	Subject	Grade
HuongNTQ	00345	MAS291	8.0
TrungTT	00786	DSA1	3.5
HienPQ	00321	DSA1	8.0

## Part 2: General Algebra

Lectures 18+19+20+21: Groups

# Groups

## **Topics covered:**

- ▶ Groups: Basic definitions and properties
- ▶ Generators for a group and cyclic groups
- ▶ Cosets and Lagrange's theorem
- ▶ Homomorphism and isomorphism
- ▶ Quotient groups

## Groups: Introduction

- ▶ In general, to study a phenomenon or a process of a real world, we construct a suitable mathematical model to represent it and to study the properties of a model to understand the phenomenon. Mathematical structures considered most are group, lattice, semigroup, etc.
- ▶ Group structure are applied in formal languages, coding theory, sequential error-correcting codes.

# Groups

Let  $G$  be a non-empty set.

- A binary operation  $*$  on  $G$  is a map

$$*: G \times G \rightarrow G$$

$$(a, b) \mapsto a * b$$

- A group is a pair  $(G, *)$  of set  $G$  together with a binary operation  $*$  on  $G$  that satisfies the following properties

1.  $G$  is closed with respect to  $*$ :

$$a + b \in G, \quad \text{for all } a, b \in G.$$

2.  $*$  is associative, i.e.,  $a * b * c = a * (b * c)$

3.  $G$  has an element  $e$ ,  $e$  is called identity element of  $G$ , such that

$$e * a = a * e = a, \quad \text{for all } a \in G.$$

4. For each  $a \in G$ , there exist an element  $b$  such that

$$a * b = b * a = e.$$

Then  $b$  is called the inverse of  $a$ , and denoted by  $a^{-1}$ .

## Examples I

- ▶ The set of integer numbers with standard addition  $(\mathbb{Z}, +)$  forms a group with the identity  $0$  and the inverse of  $a$  is  $-a$ .
- ▶ The set  $\mathbb{Z}_n = \{0, 1, \dots, n-1\}$  with the addition modulo  $n$ , i.e.,  $a+_n b = a + b \bmod n$ , forms a group, where  $e = 0$  and  $a^{-1} = \begin{cases} n - a, & \text{if } a \neq 0 \\ a, & \text{if } a = 0 \end{cases}$ .

For instance, on  $(\mathbb{Z}_7, +_7)$ , we have

$$5 +_7 6 = 5 + 6 \bmod 7 = 4 \in \mathbb{Z}_7,$$

and  $5^{-1} = 2$ .

- ▶ The set of invertible matrices of size  $n \times n$  together with matrix multiplication forms a group where the identity element is the identity matrix and the inverse of a matrix  $A$  is its inverse matrix  $A^{-1}$ .

## Examples II

- ▶ The set of non-zero rationals  $\mathbb{Q}^*$  with the standard multiplication forms a group with identity  $e = 1$  and  $\left(\frac{p}{q}\right)^{-1} = \frac{q}{p}$  for any  $p, q \in \mathbb{Z}^*$ .

## Examples III

- ▶ The set  $\mathbb{Z}_n = \{0, 1, \dots, n - 1\}$  with the multiplication modulo  $n$ , i.e.,  $a \cdot_n b = a \cdot b \bmod n$ , is not a group since there exist some elements, for instance  $0$ , not invertible by multiplication modulo  $n$ .

For example, on  $(\mathbb{Z}_6, \cdot_6)$ , we have

$$2 \cdot_6 4 = 2 \cdot 4 \bmod 6 = 2,$$

and the elements  $0, 2, 3, 4$  are not invertible (see the table of multiplication modulo  $6$  below).

## Examples IV

	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6
2	0	2	4	6	1	3	5
3	0	3	6	2	5	1	4
4	0	4	1	5	2	6	3
5	0	5	3	1	6	4	2
6	0	6	5	4	3	2	1

Table: Multiplication table on  $\mathbb{Z}_7$

	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

Table: Multiplication table on  $\mathbb{Z}_6$

## Examples V

- The set  $\mathbb{Z}_n^*$  together with the multiplication modulo  $n$ , where  $\mathbb{Z}_n^*$  is the set of all invertible elements, forms a group, denoted by  $(\mathbb{Z}_n^8, \cdot_n)$ .

For instance, the multiplicative tables for  $(\mathbb{Z}_6^*, \cdot_6)$  and  $(\mathbb{Z}_{10}^*, \cdot_{10})$  are given below, where  $\mathbb{Z}_6^* = \{1, 5\}$  and  $\mathbb{Z}_{10}^* = \{1, 3, 7, 9\}$

$\cdot_6$	1	5
1	1	5
5	5	1

Table: Group  $(\mathbb{Z}_6^*, \cdot_6)$  with  
 $5^{-1} = 5$  as  $5 \cdot_6 5 = 25$   
 $\text{mod } 6 = 1$ .

$\cdot_{10}$	1	3	7	9
1	1	3	7	9
3	3	9	1	7
7	7	1	9	3
9	9	7	3	1

Table:  $(\mathbb{Z}_{10}^*, \cdot_{10})$  with  
 $3^{-1} = 7, 7^{-1} = 3$ .

## Group $(\mathbb{Z}_n^*, \cdot_n)$

### Theorem

Let  $a \in \mathbb{Z}_n$ . Then,

- i)  $a \in \mathbb{Z}_n^*$  if and only if  $\gcd(a, n) = 1$ .
- ii) The inverse of  $a \in \mathbb{Z}_n^*$  can be found by using the extended Euclid's algorithm from the bottom to the top.

Consequently,  $\mathbb{Z}_n^*$  has  $\phi(n)$  elements, where  $\phi(n)$  is Euler's totient function of  $n$ .

**Example:** As we mentioned that  $1, 3, 7, 9$  are invertible in  $\mathbb{Z}_{10}$  with multiplication modulo  $10$  and  $6$  is not invertible since  $\gcd(6, 10) \neq 1$ .

# Power of an element in a group

Given a group  $(G, *)$ , an element  $a \in G$ , and an integer number  $s$ . We denote

$$a^s = \begin{cases} \underbrace{a * a * \cdots * a}_{s \text{ times}}, & \text{if } s > 0 \\ \underbrace{a^{-1} * a^{-1} * \cdots * a^{-1}}_{-s \text{ times}}, & \text{if } s < 0 \\ e, & \text{if } s = 0 \end{cases}$$

Then  $a^s$  is called the  $s$ th power of  $a$ .

## Examples:

- ▶ Let  $(\mathbb{Z}_6, +_6)$ . Then  $2 +_6 2 = 4$  is the second power of 2, and  $0 = 2 +_6 2 +_6 2$  is the third power of 2,  
 $3^{-1} +_6 3^{-1} = (-3) +_6 (-3) = 0$  is the  $(-2)$ nd power of 3.
- ▶ Let  $(\mathbb{Z}_{10}^*, \cdot_{10})$ . Then 1, 3, 9, 7 are 0, 1st, 2nd, 3rd powers of 3. Moreover,  $3^{-1} \cdot_{10} 3^{-1} = 7 \cdot_{10} 7 = 9$  is also the  $(-2)$  power of 3.

# Properties of a group

## Lemma

Let  $(G, *)$  be a group. The followings hold true

- (i) The identity element of  $G$  exists unique
- (ii) The inverse of an element  $a$  defined unique.
- (iii) The left (resp. right) cancellations hold, i.e., if  $ab = ac$  (resp.  $ba = ca$ ), then  $b = c$ .
- (iv)  $(ab)^{-1} = b^{-1}a^{-1}$  for any  $a, b \in G$ .
- (v)  $a^{s+t} = a^s * a^t$  and  $a^{st} = (a^s)^t$ .

## Proof.

On the whiteboard. □

# Order of a group

Let  $(G, *)$  be a group.

- The **order** of  $G$ , denoted by  $|G|$ , is the number of elements in  $G$  if it is finite, and is  $\infty$  if it is not.
- The group  $G$  is called **abelian** or **commutative** if the binary operation  $*$  is commutative, i.e.,  $a * b = b * a$  for all  $a, b \in G$ .
- A subset  $G'$  of  $G$  is called a **subgroup** of the group  $(G, *)$  if  $(G', *)$  is a group. We denote  $G' \leq G$ .

## Examples:

1. The group  $(\mathbb{Z}_n, +_n)$  is of order  $n$  since  $|\mathbb{Z}_n| = n$
2. The group  $(\mathbb{Q}^*, \cdot)$  is of order  $\infty$
3. The group  $(\mathbb{Z}_{10}^*, \cdot_{10})$  has order 4; The group  $(\mathbb{Z}_{2^k}^*, \cdot_{2^k})$  has order  $2^{k-1}$ . Generally, the group  $(\mathbb{Z}_n^*, \cdot_n)$  is of order  $\phi(n)$  where  $\phi(n)$  is the Euler totient function for  $n$ .

## Theorem

Given a group  $(G, *)$  and  $G' \subseteq G$ . Then  $G'$  is a subgroup of  $G$  if and only if for any  $a, b \in G'$  we have  $a * b^{-1} \in G'$ . Consequently, if  $|G'| < \infty$ , then  $(G', *)$  is a subgroup of  $G$  iff  $G'$  is closed under  $*$ .

## Proof.

Refer to textbooks.



## Cyclic groups

Given a group  $(G, *)$  and  $a \in G$ . Denote

$$\langle a \rangle = \{a^n : n \in \mathbb{Z}\}$$

the subgroup generated by powers of  $a$ .

**Examples:**

- ▶ In  $(\mathbb{Z}_6, +_6)$ , we have  $\langle 2 \rangle = \{0, 2, 4\}$  and  $\langle 3 \rangle = \{0, 3\}$ , and  $\langle 5 \rangle = \{0, 5, 4, 3, 2, 1\}$
- ▶ In  $(\mathbb{Z}_{12}^*, \cdot_{12})$ , we have  $\langle 5 \rangle = \{5, 1\}$  and  $\langle 7 \rangle = \{7, 1\}$ , and  $\langle 11 \rangle = \{11, 1\}$

# Cyclic groups

We have the following

## Lemma

$\langle a \rangle$  is a subgroup of  $G$ . We also say that  $\langle a \rangle$  is the subgroup of  $G$  generated by  $a$ . In addition,  $\langle a \rangle$  is called a cyclic subgroup of  $G$ .

## Definition

Order of an element Given a group  $(G, *)$  and an element  $a \in G$ .

The order of the element  $a$ , denoted by  $o(a)$ , is the order of the group  $\langle a \rangle$ .

## Example:

- ▶ In  $(\mathbb{Z}_6, +_6)$ , we have  $\langle 2 \rangle = \{0, 2, 4\}$  and  $o(2) = |\{0, 2, 4\}| = 3$ .
- ▶ In  $(\mathbb{Z}_{12}, \cdot_{12})$ , we have  $\langle 5 \rangle = \{5, 1\}$  and  $o(5) = 2$ .

# Cyclic groups

## Lemma

If  $G$  is a finite group, then the following holds true

$$o(a) = \min\{r \in \mathbb{Z}^+ : a^r = e\}.$$

**Remark:** The above lemma allows us to calculate the order of an element in a finite group as the periodic period in the orbit of its positive powers.

## Example:

- ▶ Consider sequences

$$\{2, 2 +_6 2, 2 +_6 2 +_6 2, \dots\} = \{2, 4, 0, 2, 4, 0, \dots\}$$

in  $(\mathbb{Z}_6, +_6)$ , we get  $o(2) = 3$  as showed in the previous slide.

- ▶ Consider sequences

$$\{5, 5 \cdot_{12} 5, 5 \cdot_{12} 5 \cdot_{12} 5, \dots\} = \{5, 1, 5, 1, \dots\}$$

in  $(\mathbb{Z}_{12}^*, \cdot_{12})$ , we get  $o(5) = 2$  in  $\mathbb{Z}_{12}^*$  as showed in the previous slide.

# Cyclic groups

## Definition

A group  $(G, *)$  is called **cyclic** if there exists an element  $a \in G$  such that  $G = \langle a \rangle$ .

## Examples:

- ▶  $(\mathbb{Z}, +)$  is cyclic since  $\mathbb{Z} = \langle 1 \rangle$ .
- ▶  $(\mathbb{Z}_n, +_n)$  is cyclic since  $\mathbb{Z}_n = \langle a \rangle$  for each element  $a \in \mathbb{Z}_n$  such that  $\gcd(n, a) = 1$  (please verify this!).
- ▶  $(\mathbb{Z}_{10}^*, \cdot_{10})$  is cyclic since  $\mathbb{Z}_{10}^* = \langle 3 \rangle$ .
- ▶  $(\mathbb{Z}_{12}^*, \cdot_{12})$  is not cyclic since all elements of  $\mathbb{Z}_{12}^*$  are of order 2 while  $\mathbb{Z}_{12}^*$  is of order 4.

## Direct product

Given two groups  $(G, *)$  and  $(H, \circ)$ . Consider the Cartesian product of  $G$  and  $H$ :

$$G \times H = \{(a, b) : a \in G, b \in H\}.$$

We establish a binary operation  $\diamond$  on  $G \times H$  as follows:

$$(a_1, b_1) \diamond (a_2, b_2) = (a_1 * a_2, b_1 \circ b_2), \quad \text{for all } a_1, a_2 \in G, b_1, b_2 \in H.$$

Then we have

### Lemma

$(G \times H, \diamond)$  forms a group with the identity  $(e_G, e_H)$  and the inverse  $(a, b)^{-1} = (a^{-1}, b^{-1})$ , where  $e_G, e_H$  are identities of groups  $G$  and  $H$  respectively and  $a^{-1}$  and  $b^{-1}$  are inverses of  $a$  in  $G$  and of  $b$  in  $H$  respectively.

## Examples

- In the group  $(\mathbb{Z}_3 \times \mathbb{Z}_4, \diamond)$ , we have

$$(2, 2) \diamond (1, 2) = (2 +_3 1, 2 +_4 2) = (0, 0),$$

and

$$(2, 2)^{-1} = (2^{-1}, 2^{-1}) = (1, 2),$$

where the inverse of the first 2 is taken in  $\mathbb{Z}_3$  (where  $2^{-1} = 1$  since  $1 +_3 2 = 0$  in  $\mathbb{Z}_3$ ) and the inverse of the second 2 is taken in  $\mathbb{Z}_4$ .

- In  $(\mathbb{Z}_4^* \times \mathbb{Z}_5^*, \diamond)$ , we have

$$(3, 3) \diamond (1, 4) = (3 \cdot_4 1, 3 \cdot_5 4) = (3, 2),$$

and

$$(3, 3)^{-1} = (3^{-1}, 3^{-1}) = (3, 2).$$

## Lemma

The direct product group  $\mathbb{Z}_m \times \mathbb{Z}_n$  is cyclic if and only if  $\gcd(m, n) = 1$ .

Proof.

Left for the readers.



# Cosets and Lagrange's theorem

## Definition (Cosets)

Given a group  $(G, *)$  and a subgroup  $H \leq G$ . Let  $a \in G$ . We define

$$aH = \{a * h : h \in H\}, \quad \text{and} \quad Ha = \{h * a : h \in H\}$$

left and right cosets of  $H$  in  $G$  respectively.

### Examples:

- Let  $G = \mathbb{Z}_9$  with addition modulo 9 and let  $H = \langle 3 \rangle = \{3, 6, 0\} \leq G$ . Then

$$0H = \{0 +_9 3, 0 +_9 6, 0 +_9 0\} = \{3, 6, 0\} = 3H = 6H,$$

and

$$1H = \{4, 7, 1\} = 4H = 7H, \quad 2H = \{5, 8, 2\} = 5H = 8H.$$

- Let  $G = \mathbb{Z}_{12}^*$  with the multiplication modulo 12 and let  $H = \langle 5 \rangle = \{5, 1\} \leq G$ . Then,

$$7H = \{7 \cdot_{12} 5, 7 \cdot_{12} 1\} = \{11, 7\} = 11H, \quad 1H = \{5, 1\} = 5H, \quad \text{319.}$$

## Theorem

Let  $(G, *)$  be a group and  $H \leq G$ . Then the following statements are true:

- i)  $|H| = |aH|$  for any  $a \in G$ ;
- ii) If  $b \in aH$ , then  $bH = aH$ ;
- iii) If  $b \notin aH$ , then  $aH \cap bH = \emptyset$ .

## Sketch of proof.

- i) Show that the following map is bijective:

$$\begin{aligned}\psi : H &\rightarrow aH, \\ h &\mapsto a * h.\end{aligned}$$

- ii) Prove two containments  $aH \subseteq bH$  and  $bH \subseteq aH$ .
- iii) On contrary, if  $c \in aH \cap bH$  then by (i) we get  $aH = cH = bH$  which is a contradiction.



## Index of a subgroup

**Remark:** The previous theorem shows that two cosets are either equal or disjoint. Hence, we can define the index of a subgroup  $H$  in  $G$  as follows:

### Definition

Given a group  $(G, *)$  and  $H \leq G$ . The number of left (right) cosets of  $H$  in  $G$  is called the **index** of  $H$  in  $G$ , and it is denoted by  $|G : H|$ .

### Examples:

- Let  $G = \mathbb{Z}_9$  and  $H = \langle 3 \rangle$ . Then,  $|G : H| = 3$  since there are 3 disjoint cosets of  $H$  in  $G$ :

$$0H = 3H = 6H; 1H = H = 7H; 2H = 5H = 8H.$$

- Let  $G = \mathbb{Z}_{12}^*$  and  $H = \langle 5 \rangle$ . Then,  $|G : H| = 2$  since there are 2 disjoint cosets of  $H$  in  $G$ :

$$1H = 5H; 7H = 11H.$$

# Lagrange's theorem

## Theorem

If  $G$  is a finite group and  $H$  is a subgroup of  $G$ , then  $|H|$  divides  $|G|$ . Precisely, we have

$$|G| = |H| \cdot |G : H|,$$

i.e., the number of distinct left (right) cosets of  $H$  in  $G$  is  $|G|/|H|$ .

## Proof.

The results is a corollary of the equalities or disjointness of two cosets. Details are left for readers. □

## Examples:

- ▶ A set of 3 elements cannot be a subgroup of  $\mathbb{Z}_8$ .
- ▶ The group  $\mathbb{Z}_{15}$  cannot contain a subgroup of order 6.
- ▶ Let  $G = \mathbb{Z}_9$  and  $H = \langle 3 \rangle$ . Then  $|G| = 9$ ,  $|H| = 3$  and  $|G : H| = |G|/|H| = 3$ .

# Permutations and symmetric groups

## Definition

Let  $S$  be a set of  $n$  objects  $\{1, 2, \dots, n\}$ . A permutation  $\sigma$  on  $S$  is a bijection from  $S$  to itself.

## Examples:

- ▶ Let  $S = \{1, 2, 3\}$ . Then  $\sigma(1) = 2$ ,  $\sigma(2) = 3$ ,  $\sigma(3) = 1$  is a permutation on  $\{1, 2, 3\}$ .
- ▶ All permutations on  $\{1, 2\}$  are

$$\sigma_2 : \{1, 2\} \rightarrow \{1, 2\};$$

$$\begin{array}{ll} \sigma_1 : \{1, 2\} \rightarrow \{1, 2\}; & \\ & 1 \mapsto 2, \\ & 1 \mapsto 1, \\ & 2 \mapsto 2, \end{array}$$

- ▶ There are  $n!$  permutations on a set of  $n$  objects  $\{1, 2, \dots, n\}$ .

**Exercises:** Write down all possible permutations of [4] that maps 1 to 3.

# Permutation notations I

Let  $S = \{1, 2, \dots, n\}$  and  $\sigma$  be a permutation on  $S$ .

1. One-line notation:  $\sigma(1) \ \sigma(2) \ \cdots \ \sigma(n)$ : We write all images of elements from  $1, 2, \dots, n$  of  $\sigma$  on a line
2. Two-line notation:

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

where the first line represents all elements of  $S$ , the second line represents their corresponding images.

## Permutation notations II

3. **Cycle notation:** We will decompose  $S$  into disjoint cycles to represent permutations. Each cycle is of the form  $(i_1, i_2, \dots, i_k)$  such that  $i_2 = \sigma(i_1)$ ,  $i_3 = \sigma(i_2)$ ,  $\dots$ ,  $i_k = \sigma(i_{k-1})$  and  $i_1 = \sigma(i_k)$ . Precisely,
- ▶ Starting from some element  $i$  of  $S$ , we write the sequence  $(i, \sigma(i), \sigma(\sigma(i)), \dots)$  of successive images under  $\sigma$ , until the image returns to  $i$ , at which point one closes the parenthesis rather than repeat  $i$ .
  - ▶ We then continue by choosing a new element  $j \in S$  outside the previous cycles and writing down the cycle starting at  $j$ ; and so on until all elements of  $S$  are written in cycles.

## Examples

Permutation  $\sigma(1) = 2, \sigma(2) = 5, \sigma(3) = 4, \sigma(4) = 3, \sigma(5) = 1$  is represented as

1. 25431 by one-line notation;
2.  $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix}$  by two-line notation;
3.  $(1, 2, 5)(3, 4)$  by cycle notation since
  - ▶  $1 \xrightarrow{\sigma} 2 \xrightarrow{\sigma} 5 \xrightarrow{\sigma} 1$  we get the cycle  $(1, 2, 5)$ ;
  - ▶  $3 \xrightarrow{\sigma} 4 \xrightarrow{\sigma} 3$  we get the cycle  $(3, 4)$ .

## Inverse and powers of a permutation

**Remark:** Since a permutation on  $[n]$  is in fact a bijection from  $[n]$  to  $[n]$ , we can define the inverse permutation  $\sigma^{-1}$  of  $\sigma$  as the inverse function of  $\sigma$ , and  $\sigma^k = \sigma \circ \sigma \circ \cdots \circ \sigma$  as the composite function of  $k$  times of  $\sigma$ .

**Example:** Let  $\sigma = 25431$ . Then,

$$\sigma^{-1} = 51432,$$

where  $\sigma^{-1}(1) = 5$  as  $\sigma(5) = 1$ ;  $\sigma^{-1}(2) = 1$  as  $\sigma^1 = 2$ ; etc.

$$\sigma^2 = 51342,$$

as  $\sigma(\sigma(1)) = \sigma(2) = 5$ ;  $\sigma(\sigma(2)) = \sigma(5) = 1$ ;  $\sigma(\sigma(3)) = \sigma(4) = 3$ ; etc.

## Remark

In cycle notation, it is much easier to find inverse or powers of a permutation. More precisely,

- ▶ Each cycle of  $\sigma^{-1}$  is the reflection of each cycle of  $\sigma$ .
- ▶ Cycles of  $\sigma^k$  are determined by  $k$ -step transitions in cycles of  $\sigma$ .

**Example:** Let  $\sigma = (125)(34)$  be a permutation in cycle notation. Then,

- ▶

$$\sigma^{-1} = (152)(34),$$

as the reflections of cycles 125 and 34 are 521 and 43 respectively.

- ▶

$$\sigma^2 = (1, 5, 2)(3)(4),$$

as the 2-step transition of 1, 5, 2 are 5, 2, and 1 respectively and the 2-step transition of 3 is 3, the 2-step transition of 4 is 4.

# Symmetric groups

Denote  $S_n$  by the set of all permutations on  $\{1, 2, \dots, n\}$ . Consider composition operation  $\circ$  on  $S_n$ . Then,  $\circ$  is a binary operation on  $S_n$ . Moreover, we have the following

## Lemma

The set  $S_n$  together with  $\circ$  forms a group. This group is called the symmetric group on  $\{1, 2, \dots, n\}$ .

## Proof.

- ▶ Identity element: identity bijection:  $\sigma(i) = i$  for all  $i = 1, 2, \dots, n$ .
- ▶ Inverse element is the inverse function of  $\sigma$ .



# Order of permutations in symmetric groups

## Theorem

Let  $\sigma \in S_n$  and assume that  $\sigma = c_1 c_2 \dots c_k$  is the cycle representation of  $\sigma$ . Then, the order of  $\sigma$  is determined by

$$o(\sigma) = \text{lcm}(|c_1|, |c_2|, \dots, |c_k|).$$

In other words, the order of  $\sigma$  can be calculated in terms of the least common multiple of cycle sizes of  $\sigma$ .

## Proof.

On the whiteboard. □

**Example:** Given  $\sigma = 215463 \in S_6$ .

- ▶ Cycle representation:  $\sigma = (12)(356)(4)$
- ▶  $\sigma^{-1} = (21)(653)(4)$ ,  $\sigma^2 = (1)(2)(365)(4)$
- ▶  $o(\sigma) = \text{lcm}(2, 3, 1) = 6 = o(\sigma^{-1})$ , and  
 $o(\sigma^2) = \text{lcm}(1, 1, 3, 1) = 3$ .

# Group homomorphisms and isomorphisms

## Definition

Let  $(G, *)$  and  $(H, \circ)$  be two groups.

- ▶ A mapping  $f : G \rightarrow H$  is said to be a group homomorphism if  $f(a * b) = f(a) \circ f(b)$ .
- ▶ A group homomorphism  $f$  is called an (group) isomorphism if  $f$  is a bijection. In this case, we write  $G \cong H$ .

**Remark:** Isomorphic groups have the same structure. That is if  $G \cong H$ , then they share all group theoretical properties, for instance, abelian, cyclic, order, etc.

## Group homomorphism properties

### Lemma

Let  $f : G \rightarrow H$  be a group homomorphism. Then

- ▶  $f$  maps the identity  $e_G$  of  $G$  to the identity  $e_H$  of  $H$ .
- ▶  $f$  maps the inverse of  $a \in G$  to the inverse of  $f(a)$  in  $H$ , i.e.,  $f(a^{-1}) = (f(a))^{-1}$ .
- ▶  $f(a^n) = [f(a)]^n$  for any  $n \in \mathbb{Z}$ .
- ▶ If  $a * b = b * a$ , then  $f(a) \circ f(b) = f(b) \circ f(a)$

### Lemma

Let  $f$  be a group isomorphism from  $G$  to  $H$ . Then,

- ▶ Let  $a \in G$ . Then  $o(a) = o(f(a))$ , i.e., a group isomorphism preserves the orders of its elements.
- ▶  $G$  is cyclic if and only if  $H$  is cyclic. More precisely, if  $G = \langle a \rangle$ , then  $H = \langle f(a) \rangle$ .

## Kernel and range of a group homomorphism

Given a group homomorphism  $\phi$  from  $G$  to  $H$ . We define two sets

$$\ker(\phi) = \{a \in G : \phi(a) = e_G\},$$

and

$$Im(\phi) = \{\phi(a) : a \in G\}$$

the **kernel** and **range** respectively of  $\phi$ .

### Theorem

Let  $\phi$  be a group homomorphism from  $G$  to  $H$ . Then,

- (i)  $\ker(\phi)$  is a subgroup of  $G$ ;
- (ii)  $Im(\phi)$  is a subgroup of  $H$ .

### Proof.

On the whiteboard. □

## Examples

1. Consider the set  $GL(2, \mathbb{R}) = \{A \in M_{n \times n} : \det(A) \neq 0\}$  with the matrix multiplication. Let  $\mathbb{R}^*$  be the set of non-zero real number with standard multiplication. Then,

$$\begin{aligned}\phi_1 : GL(2, \mathbb{R}) &\rightarrow \mathbb{R}^* \\ A &\mapsto \det(A)\end{aligned}$$

is a group homomorphism.

2. The map

$$\begin{aligned}\psi : (\mathbb{Z}, +) &\rightarrow (\mathbb{Z}_n, +_n) \\ a &\mapsto a \mod n\end{aligned}$$

is a group homomorphism

- 3.

$$\begin{aligned}f : (\mathbb{Z}_{12}, +_{12}) &\rightarrow (\mathbb{Z}_{12}, +_{12}) \\ x &\mapsto 3x\end{aligned}$$

is a group homomorphism.

# Normal subgroups

## Definition

Normal subgroups A subgroup  $H$  of a group  $(G, *)$  is called a normal subgroup of  $G$  if  $aH = Ha$  for any  $a \in G$ . We write  $H \triangleleft G$  to refer  $H$  is a normal subgroup of  $G$ .

## Examples:

- ▶ Every subgroup of an abelian group is normal.
- ▶  $SL(2, \mathbb{R}) = \{A \in M_{2 \times 2} : \det(A) = 1\}$  is a normal group of  $GL(2, \mathbb{R}) = \{A \in M_{2 \times 2} : \det(A) \neq 0\}$ .
- ▶  $A = \{(1)(2)(3), (123), (132)\} \triangleleft S_3$ .

## Normal subgroups

### Lemma

Let  $(G, *)$  be a group. A subgroup  $H$  of  $G$  is normal if and only if  $x * H * x^{-1} \in H$  for any  $x \in G$ .

## Quotient groups

Let  $H \triangleleft G$  be a normal subgroup of  $(G, *)$ . We denote

$$\begin{aligned} G/H &= \{\text{cosets of } H \text{ in } G\} \\ &= \{aH : a \in G\} \end{aligned}$$

and define a binary operation  $\circ$  on  $G/H$  as follows

$$\begin{aligned} \circ : G/H \times G/H &\rightarrow G/H \\ (aH, bH) &\mapsto (a * b)H \end{aligned}$$

We have the following

Lemma

$(G/H, \circ)$  forms a group.

Proof.

On the whiteboard.



## Examples I

1. Consider the group  $(\mathbb{Z}, +)$  and a subgroup  $H = \langle 4 \rangle$  of it.  
Then,  $H = \{0, \pm 4, \pm 8, \dots\} \triangleleft \mathbb{Z}$ .

- ▶ Cosets of  $H$  in  $\mathbb{Z}$  are

$$[0] := 0 + H = \{0, \pm 4, \pm 8, \dots\} = \pm 4 + H = \pm 8 + H = \dots$$

$$[1] := 1 + H = \{\dots, -7, -3, 1, 5, 9, \dots\} = -7 + H = -3 + H = \dots$$

$$[2] := 2 + H = \{\dots, -6, -2, 2, 6, 10, \dots\} = -2 + H = 2 + H = \dots$$

$$[3] := 3 + H = \{\dots, -5, -1, 3, 7, 11, \dots\} = -5 + H = 3 + H = \dots$$

- ▶ The quotient group:

$$\mathbb{Z}/\langle 4 \rangle = \mathbb{Z}/4\mathbb{Z} = \{0+H, 1+H, 2+H, 3+H\} = \{[0], [1], [2], [3]\}.$$

- ▶ Binary operation on  $G/H$ :

$$[2] \circ [3] = (2 + H) \circ (3 + H) = (2 + 3 + H) = 1 + H = [1]$$

- ▶ It is the fact that

$$\mathbb{Z}/4\mathbb{Z} \cong \mathbb{Z}_4.$$

## Examples II

2. Consider the group  $(\mathbb{Z}_{18}, +_{18})$  and  $H = \langle 6 \rangle = \{0, 6, 12\}$ . Then

- ▶  $H$  is a normal subgroup of  $G$ :  $H \triangleleft \mathbb{Z}_{18}$
- ▶ Cosets of  $H$  in  $G$  are  $[0] = 0 + H$ ,  
 $[1] = 1 + H, \dots, [5] = 5 + H$ .
- ▶ Quotient group  
 $G/\langle 6 \rangle = \{0 + H, 1 + H, 2 + H, 3 + H, 4 + H, 5 + H\}$
- ▶ Binary operation on  $G/\langle 6 \rangle$ :  
 $[3] \circ [5] = (3 + H) \circ (5 + H) = 8 + H = 2 + H = [2]$ .
- ▶ It is proved that  $\mathbb{Z}_{18}/H \cong \mathbb{Z}_6$

## Examples III

3. Consider  $(\mathbb{Z}_{18}^*, \cdot_{18})$  and  $H = \{1, 7, 13\}$ . Then

- ▶  $H$  is normal in  $\mathbb{Z}_{18}^*$ :  $H \triangleleft \mathbb{Z}_{18}^*$
- ▶ Cosets of  $H$  in  $\mathbb{Z}_{18}^*$ :

$$[1] = 1H = 7H = 13H,$$

and

$$[5] = 5H = 11H = 17H.$$

- ▶ Quotient group  $G/H = \{[1], [5]\}$ .

## Remark:

From Lagrange's theorem, if  $|G| < \infty$ , then the order of quotient group  $G/H$  is equal to  $|G|/|H|$ . For instance,

- ▶ in Example 2,

$$|\mathbb{Z}_{18}/\langle 6 \rangle| = |\mathbb{Z}_{18} : H| = \frac{|\mathbb{Z}_{18}|}{|\langle 6 \rangle|} = \frac{18}{3} = 6.$$

- ▶ in Example 3,

$$|\mathbb{Z}_{18}^*/\{1, 7, 13\}| = \frac{|\mathbb{Z}_{18}^*|}{|\{1, 7, 13\}|} = 2.$$

Furthermore,  $\mathbb{Z}_{18}^*/\{1, 7, 13\} \cong \mathbb{Z}_2$  (Exercise!)

# The first group isomorphism theorem

## Theorem

Given two groups  $(G_1, *)$  and  $(G_2, \circ)$ . Let  $\phi : G_1 \rightarrow G_2$  be a group homomorphism from  $G_1$  to  $G_2$ . Then,

- ▶  $\ker(\phi)$  is a normal subgroup of  $G_1$ .
- ▶  $\phi$  induces the following group isomorphism

$$\begin{aligned}\bar{\phi} : G_1 / \ker(\phi) &\rightarrow \text{Im}(\phi) \\ x * \ker(\phi) &\mapsto \phi(x)\end{aligned}$$

## Proof.

On the whiteboard. □

## Examples

Consider the map

$$\begin{aligned}\phi : \mathbb{Z} &\rightarrow \mathbb{Z}_4 \\ x &\mapsto x \mod 4\end{aligned}$$

We can show that

- $\phi$  is a group homomorphism and is a surjection.
- $\ker(\phi) = \langle 4 \rangle$ . Consequently,  $\mathbb{Z}/\langle 4 \rangle \cong \mathbb{Z}_4$  as mentioned.