Northeastern University, Department of Mathematics

MATH 4570 - Matrix Methods in Data Analysis and Machine Learning

• Instructor: **He Wang** Email: **he.wang@northeastern.edu**

§1. Fields

1. Background:

Definition 1. (1) A **linear equation** in variables $x_1, x_2, \dots, x_n$ is of the form

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

Here, $a_1, a_2, \dots, a_n \in \mathbb{R}$ (or a field $\mathbb{F}$) are **coefficients**.

(2) A **system of linear equations** (or **linear system**) is a collection of linear equations in the same variables.

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

Matrix/vector notation:

$$A = [a_{ij}] = \begin{array}{cccc} \vec{a}_1 & \vec{a}_2 & & \vec{a}_n \\ \| & & & \end{array}$$

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}_{m\times n} \qquad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m$$

- $[A \mid \vec{b}]$
- $x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{b}$
- $A\vec{x} = \vec{b}$

Goal: Find the set of all solutions.

Method: Gauss-Jordan elimination (Gaussian elimination).

> **Theorem 2.** *A linear system (matrix equation $A\vec{x} = \vec{b}$) has either no solution, or exactly one solution, or infinitely many solutions.*

2. Sets and functions

> **Definition 3.** A **set** S is a "*well-defined*", *unordered* collection of "*distinct*" elements.

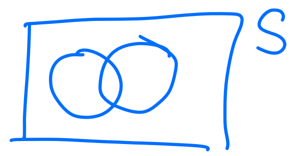
Non-well-defined example, (Russell' s paradox):

$S \in S \Leftrightarrow S \notin S$

$\boxed{S = \{x \mid x \notin x\}}$, i.e., set of all sets that are not members of themselves.

The teacher that teaches all who don't teach themselves.

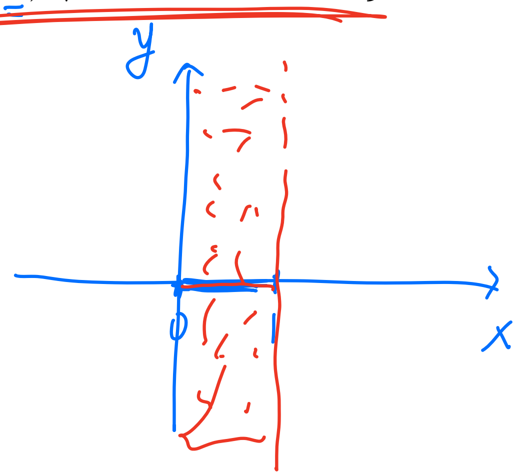
Review of set operations: A, B subsets of S



- **Union** $A \cup B = \{x \mid x \in A \text{ or } x \in B\}$
- **Intersection** $A \cap B = \{x \mid x \in A \text{ and } x \in B\}$
- **Complement** of $A \subset S$, $A^c = \{x \in S \mid x \notin A\}$
- **(Cartesian) Product** $\boxed{X \times Y} = \{(x, y) \mid x \in X, y \in Y\}$.

Ex. $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$

(x, y)



Ex: $[0, 1] \times \mathbb{R}$

$\{(x, y) \mid 0 \leq x \leq 1\}$

Ex: $f: \mathbb{R} \to \{0, 1\}$

$g: \mathbb{R}^d \to \{1, 2, \dots$

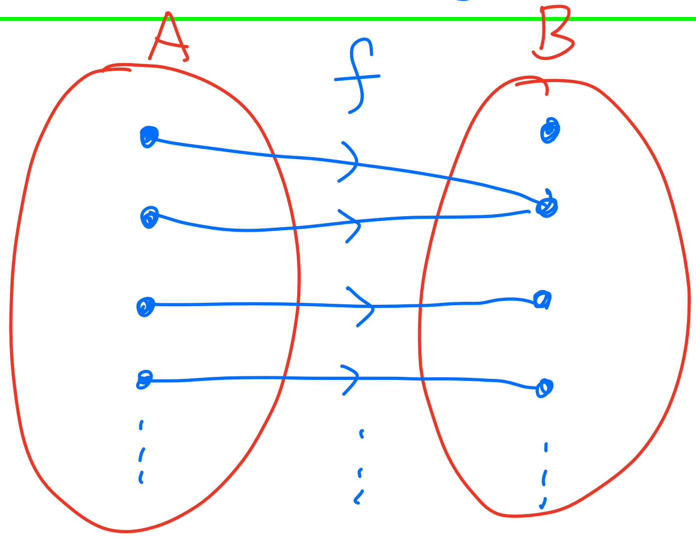
$f: \mathbb{R} \to \mathbb{R}$

$y = x + 1$ $y = e^x$ $f(x) = e^x$

Definition 4. A **function**(map) f between two sets A and B is a rule

$$f : A \to B$$

sending every $a \in A$ to an element $f(a) \in B$.

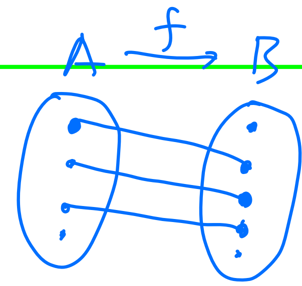


one to one ✓

two to one ✓

one to more ✗

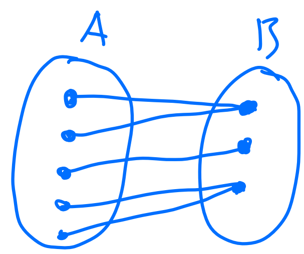
Definition 5. Let $f : A \to B$ be a function.



(1) f is called **injective (one-to-one)**, if

$f(x)=f(y)$ implies $x=y$ for any $x,y \in A$.

(2) f is called **surjective (onto)**, if

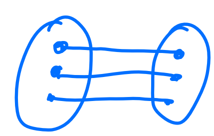


$\forall\, b \in B,\ \exists\, x \in A,$ s.t. $f(x)=b$.

for any — there exists — such that

(3) f is called **bijective**, if

f is surjective and injective.



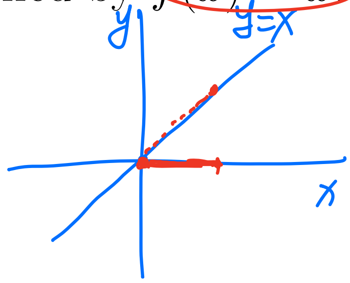
Consider a function $f : A \to B$ and the equation $f(x) = b$ for every $b \in B$.

Proposition 6.
- *f is injective $\Leftrightarrow$ $f(x) = b$ has* at most one *solution.*
- *f is surjective $\Leftrightarrow$ $f(x) = b$ has* at least one *solution.*
- *f is bijective $\Leftrightarrow$ $f(x) = b$ has* exactly one *solution.*

Example 7. Consider functions

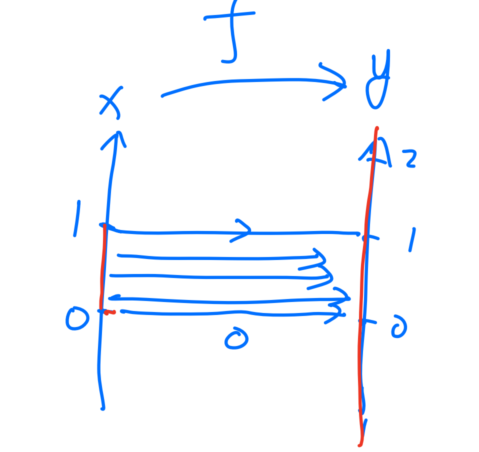
$f : [0, 1] \to \mathbb{R}$ defined by $f(x) = x$.

A B



injective

not surjective

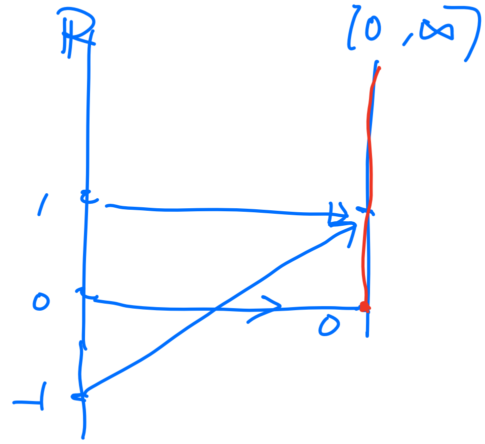


$g : \mathbb{R} \to [0, \infty)$ defined by $g(x) = x^2$.

B

surjective

not injective



$h : \mathbb{R} \to \mathbb{R}$ defined by $h(x) = 2x + 1$.

bijective.

ex. $S(x) = x^2$

$T(u) = e^u$

$T \circ S(x) = e^{x^2}$

$$x \longrightarrow S(x) \xrightarrow{T} T(S(x))$$

$$U \xrightarrow{S} V \xrightarrow{T} W$$

Definition 8. The **composition** $T \circ S(x)$ of two functions $S : U \to V$ and $T : V \to W$

$$T \circ S : U \longrightarrow W$$

$$x \longrightarrow T(S(x))$$

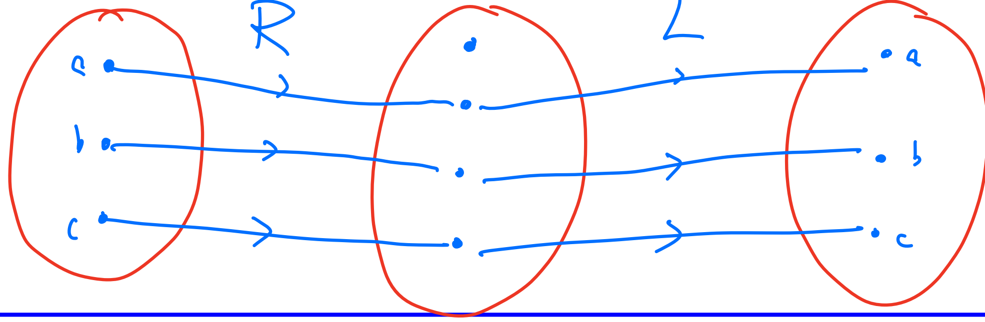
Theorem 9. *Consider functions $R : V \to W$ and $L : W \to V$. If*

left inverse right inverse

$$L \circ R = \mathrm{id}_V$$

then L is surjective and R is injective. (That is $V \overset{R}{\hookrightarrow} W \overset{L}{\twoheadrightarrow} V$)

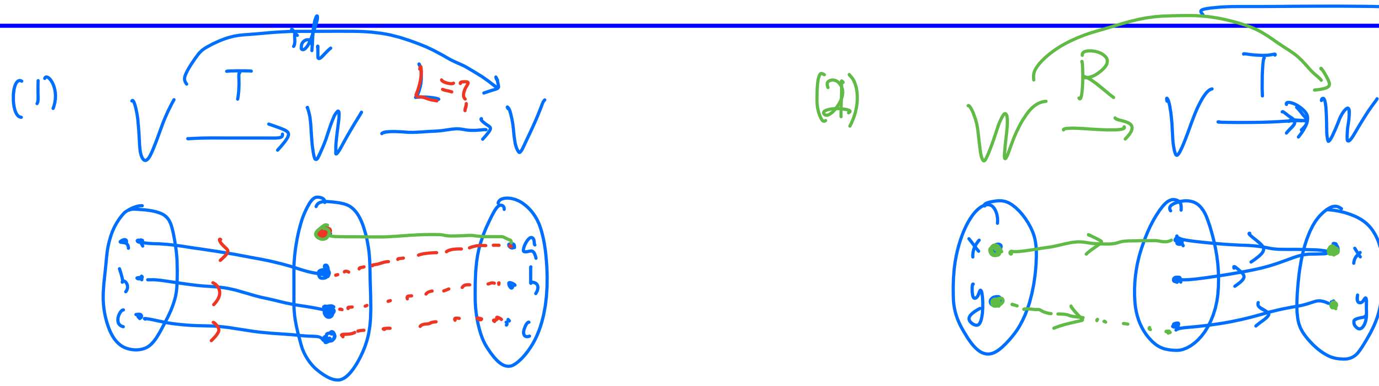
$$\begin{tikzcd} V \arrow[r, hook, "R"] \arrow[rr, bend left, "\mathrm{id}_V"] & W \arrow[r, two heads, "L"] & V \end{tikzcd}$$



Theorem 10.

$L \circ T = id_V$ not unique

(1) *A map $T : V \to W$ is injective if and only if it has a left-inverse.*
(2) *A map $T : V \to W$ is surjective if and only if it has a right-inverse.*



the inverse of T

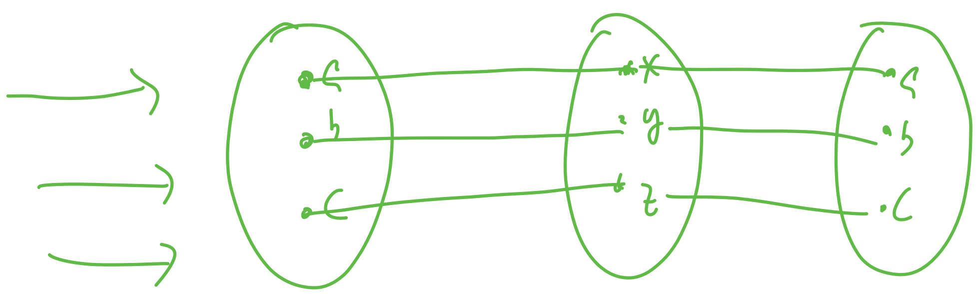
L=R

map injective

Theorem 11. *Suppose a function $T : V \to W$ has both a left-inverse L and a right-inverse R. Then*

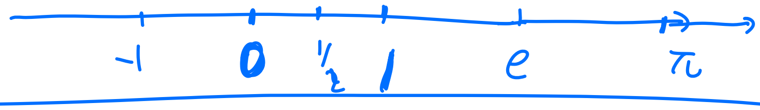
⇔ surjective

$$L = R : W \to V$$



Proposition 12. *A map $T : V \to W$ is bijective if and only if it is invertible.*

set

Ex1 $\mathbb{R} = \{$  $\}$

Two operations:

	identities	inverse of a
1. sum(+)	0	$-a$
2. product. (×)	1	$1/a$ if $a \neq 0$

$(a+b\sqrt{-1})\times(c+b\sqrt{-1})$
$\parallel$
$(\ ?\ ) + (\ ?\ )\sqrt{-1}$

$(-a) + (-b)\sqrt{-1}$

$\dfrac{1}{a+b\sqrt{-1}} = (\ ?\ ) + (\ ?\ )\sqrt{-1}$

with properties:

(1) $(a+b)+c = a+(b+c)$

(2) $a+b = b+a$

(3) $(a\times b)\times c = a\times(b\times c)$

(4) $ab = ba$

$(a+b)\times c = a\times c + b\times c$

Ex2 Set $S = \{[0], [1]\} = \mathbb{Z}_2$

even odd

$-[1] = ?$

$[1] + ? = [0]$

Two operations on S

sum +	[0]	[1]
[0]	[0]	[1]
[1]	[1]	[0]

product ×	[0]	[1]
[0]	[0]	[0]
[1]	[0]	[1]

	identities	inverses
$+$	$[0]$	notation $-[1] = [1]$
$\times$	$[1]$	$[1]^{-1} = [1]$

$[1] + (-[1]) = [0]$

$$\mathbb{C} = \{ a + bi \mid a, b \in \mathbb{R} \}$$

3. Algebraic objects:

Set → **Monoid** → **Group**→ **Ring**→ **Field**

$+$ $\times$

e.g. sum, or product

$a+b$

$a\times b$

Definition 13. A **binary operation** on a set S is a map/function

$$* : S\times S \longrightarrow S$$
$$(a, b) \mapsto a*b$$

Definition 14. A **monoid** is a set M with a binary operation $* : M\times M \to M$ satisfying two axioms:

(1) (Identity) e There exists $\exists$ $e\in M$ such that $e*x = x = x*e$ for any $x\in M$.

(2) (Associativity) $(a*b)*c = a*(b*c)$. for any $a, b, c\in M$

Proposition 15. *Identity is unique in a monoid.*

Definition 16. A monoid $(M,*)$ is called a **commutative** (or abelian), if

$a*b = b*a$ for any $a, b\in M$

Definition 17. A **group** is a monoid $(G, \cdot)$ satisfies

(3) (Inverse) For any $(\forall)$ $g \in G$, there exist $h \in G$ such that $g * h = h * g = e$

h is called the inverse of g, denoted by $h =: g^{-1}$

Proposition 18. *In a group G, inverse is unique in for any $g \in G$.*

Denote commutative (abelian) group as $(G, \underset{*}{+}, \underset{e}{0})$; inverse of a as $-a$.

Definition 19. A **ring** (with unit/identity) is a set R with two binary operations $+$ and $\cdot$, s.t.

(1) $(R, +)$ is an "abelian group."

(2) (multiplicative identity) e' $\quad e' \cdot a = a \cdot e' = a$

(3) (multiplicative associative) $\quad (a \cdot b) \cdot c = a \cdot (b \cdot c)$

(4) (Distributivity) $\quad a \cdot (b + c) = a \cdot b + a \cdot c$

$(b + c) \cdot a = b \cdot a + c \cdot a$

Definition 20. A ring R is called a **commutative** if

$$ab = b \cdot a$$

(Denote e' as 1 in commutative ring.)

Example 21. Integers $\mathbb{Z}$ is a commutative ring.

$+ \quad 0 \quad a+(-a)=0$

$\times \quad 1 \quad ?$

$2 \cdot \frac{1}{2} = 1$

Example 22. Set of all polynomials $\mathbb{R}[t]$ with sum and product is a commutative ring.

$a_0+a_1t+a_2t^2+\cdots+a_nt^n$

$a_0 \cdots a_n \in \mathbb{R}$

Example 23. Set of all polynomials $\mathbb{R}[x_1, x_2, \ldots, x_n]$ is a commutative ring.

$= \{0, \pm 2, \pm 4, \pm 6, \ldots\}$

Example 24. $2\mathbb{Z}$ is a ring without identity.

$\mathbb{Z}_2 = \{[0], [1]\}$

$\mathbb{Z} = \{\ldots -5\ -4\ -3\ -2\ -1\ 0\ 1\ 2\ 3\ 4\ 5\ 6 \ldots\}$

Definition 25. A **field** $\mathbb{F}$ is a commutative ring $(\mathbb{F}, +, \cdot)$ such that
(5) any non-zero element has a multiplicative inverse.

$0 \neq g \in \mathbb{F}$ $\qquad g^{-1} \cdot g = 1$

Remark: $(F - \{0\}, \cdot)$ are abelian groups.

Ex: $\mathbb{Z}_4 = \{[0], [1], [2], [3]\}$ ($n=4$)

$[-8] = [0] = \{0, \pm 4, \pm 8, \pm 12, \ldots, \pm 4k, \ldots\}$

$[1] = \{1, 1\pm 4, 1\pm 8, 1\pm 12, \ldots, 1\pm 4k, \ldots\}$

$[2] = \{2, 2\pm 4, 2\pm 8, \ldots\}$

$[3] = \{3, 3\pm 4, 3\pm 8, \ldots\}$

Two operations: $[a]+[b] := [a+b]$

$[a]\times[b] := [a\times b]$

- identities for sum: $[0]$ since $[0]+[a]=[a]$
- ——— prod: $[1]$ since $[1]\times[a]=[a]$

- sum inverse of $[a]$ $\quad -[a] =$
- product inverse of $[a]$ $(a \neq 0)$ $\quad \frac{1}{[a]} = [a]^{-1} =$

e.g. $-[2] = [?]$ since $[2]+[?]=[0]$

$[2]^{-1}$ = not exist since $[2]\times[?]=[1]$

$[x]=[?]=[3]^{-1}=[3]$ $\quad [3]\cdot[?]=[1]$

$[3x]=[1]$

For $n > 0 \in \mathbb{Z}$, let $\mathbb{Z}_n = \{[0], [1], \ldots, [n-1]\}$=the set of congruence classes modulo n.

Proposition 26. *$(\mathbb{Z}_n, +, \times)$ is a commutative ring.*

Example 27. $\mathbb{Z}_2$ is a field.

$\mathbb{Z}_4$

Example 28. $\mathbb{Z}_6$ is not a field. (Reason: [2] has no multiplicative inverse.)

Proposition 29. *$\mathbb{Z}_n$ is a field if and only if $n = p$ is a prime number.*

$\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are fields.

Remark: $\mathbb{Q}$ is the smallest field containing $\mathbb{Z}$.

In our class, we will focus on fields $\mathbb{R}$, $\mathbb{C}$, (and $\mathbb{Z}_p$).

$\mathbb{Z}_2$

The idea of group and field was created by Évariste Galois (1811 – 1832).

$ax^2+bx+c=0$

$a_1x^4+a_2x^3+a_3x^2+a_4x+a_5=0$

https://en.wikipedia.org/wiki/%C3%89variste_Galois

$A \xrightarrow{f} B$

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

map

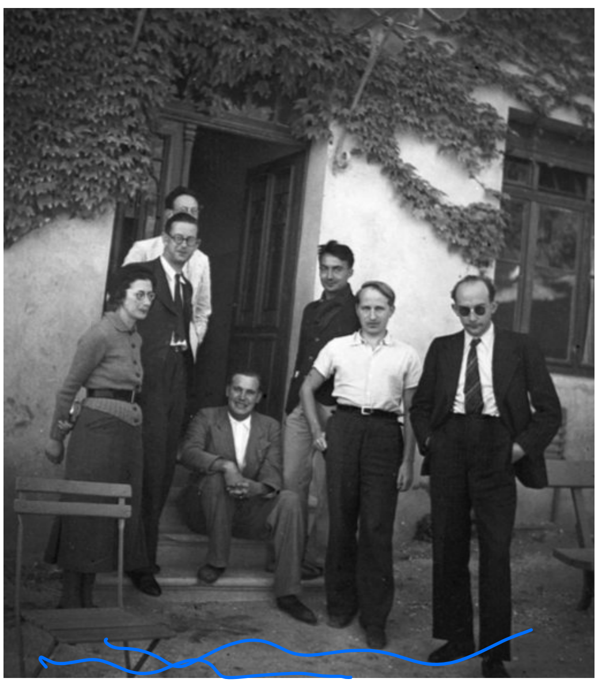
4. Functions between algebraic objects:

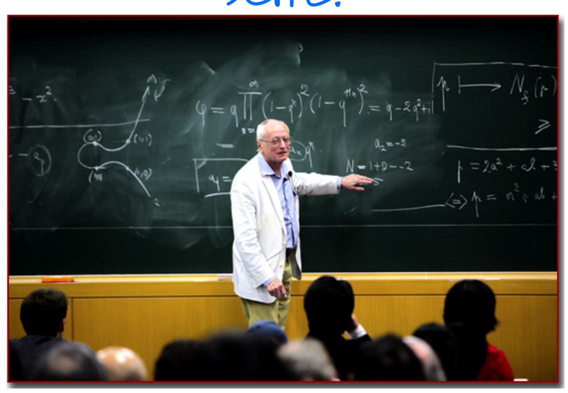
Definition 30. A **homomorphism** $f : A \to B$ between any two algebraic objects is a function preserving all operations, i.e.,

$$f(x * y) = f(x) * f(y) \text{ for any } x, y \in A$$

For ring with identity, we also need the homomorphism sends identity to identity.

Definition 31. (Terminology first by Nicolas Bourbaki (1934-).)
(1) An injective homomorphism is called **monomorphism**.
(2) A surjective homomorphism is called an **epimorphism**.
(3) A function $f : A \to B$ is called **isomorphism**, if it is monomorphism and epimorphism. In this case, we consider A and B are the "same".



Serre.


https://en.wikipedia.org/wiki/Nicolas_Bourbaki

Further extended reading:

1. Classification finite fields.
2. Classification of finite abelian groups.
3. "Classification of finite groups".