Northeastern University, Department of Mathematics

Linear Algebra

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

§1. Linear system and Gaussian elimination over fields

Topics: 1. Linear system; 2. Sets, groups, fields and more; 3. Gaussian elimination.

1. Background:

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

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

Here, $a_1, a_2, \ldots, a_n \in \mathbb{R}$ (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.

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

Matrix/vector notation:

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

- **Coefficient matrix** A: Size $m \times n$; m **rows**; n **columns**.
- **Vector** $\vec{b} \in \mathbb{R}^m$ (or $\mathbb{F}^n$).
- **Augmented matrix**: $[A \mid \vec{b}\,]$.

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):

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

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

Set operations:

- **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 $A \subset S$, $A^c = \{x \in S \mid x \notin A\}$
- **Product** $X \times Y = \{(x, y) \mid x \in X, y \in Y\}$.
  E.g., $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$.

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 $f(a) \in B$

It is ok to have $f(a_1) = f(a_2)$ for different a_1 and a_2. It is wrong to send one element in A to two different elements in B.

Definition 5.
(1) A function $f : A \to B$ is called **injective (one-to-one)**, if $x \neq y$ implies $f(x) \neq f(y)$, or equivalently, $f(x) = f(y)$ implies $x = y$ for any $x, y \in A$.
(2) A function $f : A \twoheadrightarrow B$ is called **surjective (onto)**, if for any $b \in B$, there is $x \in A$ such that $f(a) = b$.
(3) A function $f : A \hookrightarrow B$ is called **bijective**, if it is both injective and surjective.

Consider a function $f : A \to B$ and the equation $f(x) = b$ for every $b \in B$. From the definition, we can get the following properties.

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 $f(x) = x$.

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

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

Definition 8. The **composition** $T \circ S$ of two functions $S : U \to V$ and $T : V \to W$ is the function

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

defined by $(T \circ S)(u) = T(S(u))$ for $u \in U$.

Theorem 9. *Consider functions $R : V \to W$ and $L : W \to V$. Let id_V be the* ***identity*** *map of V defined by $id_V(v) = v$ for all $v \in V$.*
If

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

then L is surjective and R is injective. (That is $V \overset{R}{\hookrightarrow} W \overset{L}{\twoheadrightarrow} V$)
In this case, we call L a ***left-inverse of*** *R (i.e., R has a left-inverse); and call R a* ***right-inverse of*** *L.*

Proof. Directly from definitions of surjective and injective, and consider $L \circ R : V \to W \to V$.
For any $v \in V$, $L \circ R(v) = v$, so $L(R(v)) = v$, hence L is surjective.
Suppose $R(v_1) = R(v_2)$. Apply L both sides, we have $v_1 = v_2$. So, R is injective.

□

Theorem 10. *(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.

Proof. "$\Leftarrow$" is from the above theorem.
(1) "$\Rightarrow$" Since T is injective, for each $w \in W$, the equation $w = T(x)$ has at most one solution. If $w = T(x)$ has a (unique) solution x, then define a map $L : W \to V$ as $L(w) = x$. If there is no solution, we can assign any value for w. Then L is a left-inverse of T. (Notice that, it is not unique.)
(2)"$\Rightarrow$" Since T is surjective, for each $w \in W$, the equation $w = T(x)$ has at least one solution (maybe not unique). Choose one solution and define $R : W \to V$ as $R(w) = x$. Then R is a right-inverse of T. (Notice that, it is not unique.)

□

Theorem 11. *Suppose a function $T : V \to W$ has both a left-inverse (i.e., $L \circ T = \mathrm{id}_V$) and a right-inverse (i.e., $T \circ R = \mathrm{id}_W$).*
Then $L = R : W \to V$.
In this case, this unique function is called ***the inverse of T.*** *The function T is called* ***invertible***.

Proof. For any $w \in W$, $T(R(w)) = w$. So $L(w) = L(T(R(w))) = R(w)$. □

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

3. Algebraic objects: Set → Monoid → Group → Ring → Field

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

$$\begin{aligned} * : S \times S &\to S \\ (x, y) &\to x * y \end{aligned}$$

Definition 14. A **monoid** is a set M with a binary operation $* : M \times M \to M$ s.t.
(1) $\exists\, e \in M$, s.t. $e * x = x * e = x$, $\forall\, x \in M$. (Identity)
(2) $(a * b) * c = a * (b * c)$, $\forall\, a, b, c \in M$. (Associativity)

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

Proof. Suppose $\exists$ two identities e and $e' \in M$. Then
$$e' = e * e' = e.$$
□

Definition 16. A monoid $(M, *)$ is called a **commutative** (or abelian), if $\forall\, a, b \in M$,
$$a * b = b * a$$

Definition 17. A **group** is a monoid $(G, \cdot)$ satisfies one more axiom:
(3) $\forall\, g \in G$, $\exists\, h \in G$ s.t. $g \cdot h = h \cdot g = e$, (Inverse)

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

Proof. Suppose $\exists$ two inverses $\exists\, h, h' \in G$ for $g \in G$. $h' = h' \cdot e = h' \cdot (g \cdot h) = (h' \cdot g) \cdot h = e \cdot h = h$.
□

Denote commutative (abelian) group as $(G, +, 0)$; 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) $\exists\, e \in R$, s.t. $\forall a \in R$, $e \cdot a = a \cdot e = a$. (multiplicative identity)
(3) $\cdot$ is associative.
(4) $a \cdot (b + c) = a \cdot b + a \cdot c$,
$(b + c) \cdot a = b \cdot a + c \cdot a$, $\forall a, b, c \in R$. (Distributivity)

Definition 20. A ring R is called a **commutative** if $\forall a, b \in R$, $a \cdot b = b \cdot a$.

(Denote e as 1 in commutative ring.)

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

Example 22. Set of all polynomials $\mathbb{R}[x_1, x_2, \ldots, x_n]$ with sum and product is a commutative ring.

Example 23. $2\mathbb{Z}$ is a ring without identity.

Definition 24. A **field** $\mathbb{F}$ is a commutative ring $(\mathbb{F}, +, \cdot)$ satisfying

$$\forall a \neq 0 \in \mathbb{F}, \exists x \in \mathbb{F} \ s.t. \ ax = e$$

i.e., any nonzero element has a multiplicative inverse.

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

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

Proposition 25. *$(\mathbb{Z}_n, +, \times)$ is a commutative ring.*

Example 26. $\mathbb{Z}_2$ is a field.

Example 27. $\mathbb{Z}_6$ is not a field. (Reason: [2] has no multiplicative inverse.)

Proposition 28. *$\mathbb{Z}_n$ 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 $\mathbb{Z}_p$.

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

Function between algebraic objects:

Definition 29. A **homomorphism** $f : A \to B$ between any two algebraic objects is a function preserving all operations, i.e., $f(x * y) = f(x) * f(y)$ for any $x, y \in A$.

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

Definition 30. (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".
(Terminology first by Nicolas Bourbaki (1934-).)

Further extended reading: 1. Classification finite fields. 2. Classification of finite abelian groups. 3. "Classification of finite groups".

4. Gauss-Jordan Elimination

Go back to matrix $[A \mid \vec{b}]$.

The leftmost nonzero entry of a row is called **leading entry**(or **pivot**).

Definition 31. A matrix is in ***row-echelon form*** **(ref)** if
(1.) All entries in a column below a leading entry are zeros.
(2.) Each row above it contains a leading entry further to the left.
A matrix is in ***reduced row-echelon form*** **(rref)**, if it satisfies (1) (2) and
(3.) The leading entry in each nonzero row is 1.
(4.) All entries in a column above a leading entry are zeros.

Condition 2 implies that all zero rows are at the bottom of the matrix.

One example of **ref**, ($\blacksquare$: non-zero number, $*$ any number) and one example of **rref**

$$\textbf{ref} = \begin{bmatrix} \blacksquare & * & * & * & * & * \\ 0 & 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \to \cdots \to \quad \textbf{rref} = \begin{bmatrix} 1 & * & 0 & 0 & 0 & * \\ 0 & 0 & 1 & 0 & 0 & * \\ 0 & 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Elementary Row Operations:

(1.) **Scaling:** Multiply a row R_i by a nonzero scalar $k \neq 0$. (kR_i)
(2.) **Replacement:** Replace a row R_i by adding a multiple of another row kR_j. ($R_i + kR_j$)
(3.) **Interchange:** Interchange two rows. ($R_i \leftrightarrow R_j$)

Elementary row operations do not change solutions of the linear system.

Theorem 32. *Using the elementary row operations, one can change a matrix to a reduced row-echelon form.*

Proof. Gauss-Jordan elimination:
1. Begin with the *leftmost* ***nonzero*** column.
2. Select a *nonzero* entry as a **pivot**, and interchange its row to the first row.
3. Use ERO to create zeros in all positions below the pivot.
4. Omit the first row and repeat this process.
5. Repeat the process until the last nonzero row.
6. Scale all pivots to 1's.
7. Beginning with the **rightmost** pivot and working upward and to the left. □

Theorem 33. *A matrix A has a unique reduced row echelon form* ***rref***(A).

Proof. We outline a better method here. We will fill the details after we learned subspaces.
Step 1. Augmented matrices $[A\ \vec{0}]$ and $[\textbf{rref}(A)\ \vec{0}]$ have the same solution set, since elementary row operations do not change solution set.
Step 2. Different reduced row echelon forms have different solutions sets. □

Definition 34. If $A \xrightarrow{ERO} \cdots \xrightarrow{ERO} B$, then A is called **row-equivalent** to B.

Proposition 35. *Row-equivalent is an equivalent relation.*

Proof. 1. $A \sim A$.
2. If $A \sim B$, then $B \sim A$.
3. If $A \sim B$ and $B \sim C$, then $A \sim C$. □

Theorem 36. *A linear system* $[A|\vec{b}]$ *is inconsistent (no solution)* if and only if ***rref***$([A|\vec{b}])$ *has a row* $[\ 0\ 0\ 0\ \ldots\ 0 \mid 1\ ]$.
If a linear system is consistent, it has either
- *a unique solution (no free variables), or*
- *infinitely many solutions (at least one free variable).*

Definition 37. The **rank** of a matrix A is defined to be the number of pivots in **rref**(A), denoted as $\text{rank}(A)$.

Proposition 38. *Row-equivalent matrices have the same rank.*

Example 39. Suppose the coefficient matrix A is of size $m \times n$. Then,

1. $\text{rank}(A) \leq m$ and $\text{rank}(A) \leq n$.

2. If the system is inconsistent, then $\text{rank}(A) < m$.

3. If the system has exactly one solution, then $\text{rank}(A) = n$.

4. If the system has infinitely many solutions, then $\text{rank}(A) < n$.

Proof. The linear system of m equations with n variables. Use the rank A=number of pivots=n-number of free variables. □

Definition 40. An $m \times n$ matrix A has **full rank**, if $\text{rank}(A) = \min(m, n)$.

Proposition 41. *A linear system with an* $n \times n$ *coefficient matrix* A *has exactly one solution if and only if* $\text{rank}(A) = n$ *if and only if* ***rref***$(A) = I_n$.

Remark: 1. We can apply Gaussian elimination over integers $\mathbb{Z}$. However, we can not achieve **rref**.

2. Buchberger's algorithm is a generalization of Gaussian elimination to polynomials to obtain a Grobnear basis in commutative algebra.