VV285 RC Part I

Elements of Linear Algebra

"Matrices are just linear maps!"

Xingjian Zhang

Univerity of Michigan-Shanghai Jiao Tong University Joint Institute

May 4, 2020

1. Systems of Linear Equations
2. Finite-Dimensional Vector Spaces
3. Inner Product Spaces
4. Linear Maps
5. Matrices
6. Theory of Systems of Linear Equations
7. Determinants

Overview

1. Linear System
   Homogeneous vs. Inhomogeneous
   Underdetermined vs. Overdetermined
2. Equivalency of Linear System
3. The Gauß – Jordan Algorithm
4. Diagonalizable (Existence and Uniqueness of Linear System)
5. **Fundamental Lemma for Homogeneous Equations**

A *linear system* of m (algebraic) equations in n unknowns $x_1, x_2, \dots, x_n \in V$ is a set of equations

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

where $b_1, b_2, \dots, b_m \in V$ and $a_{ij} \in \mathbb{F}, i = 1, \dots, m,\ j = 1, \dots, n$.
If $b_1 = b_2 = \cdots = b_m = 0$, then (1) is called a *homogeneous system*.
Otherwise, it is called an *inhomogeneous system.*
If $m < n$ we say that the system in *underdetermined*, if $m > n$ the system is called *overdetermined*. A solution of a linear system of equations (1) is a tuple of elements $(y_1, y_2, \dots, y_n) \in V^n$ such that the predicate (1) becomes a true statement.

We say that two systems of linear equations are *equivalent* if any solution of the first system is also a solution of the second system and vice-versa. Thus the systems

$$\begin{aligned} x_1 + 3x_2 - x_3 &= 1 \\ -5x_2 + x_3 &= 1 \\ 10x_2 + x_3 &= 1 \end{aligned} \qquad \text{and} \qquad \begin{aligned} x_1 &= 2 \\ x_2 &= 0 \\ x_3 &= 1 \end{aligned}$$

are *equivalent*.

Gauß-Jordan Algorithm

The goal of the *Gauß-Jordan algorithm* (also called Gaußian elimination) is to transform a system

$$\begin{array}{ccc|c} * & * & * & \diamond \\ * & * & * & \diamond \\ * & * & * & \diamond \end{array} \qquad\qquad * \in \mathbb{R} \text{ or } \mathbb{C}, \quad \diamond \in V$$

first into the form

$$\begin{array}{ccc|c} 1 & * & * & \diamond \\ 0 & 1 & * & \diamond \\ 0 & 0 & 1 & \diamond \end{array} \tag{2}$$

and subsequently into

$$\begin{array}{ccc|c} 1 & 0 & 0 & \diamond \\ 0 & 1 & 0 & \diamond \\ 0 & 0 & 1 & \diamond \end{array} \tag{3}$$

Include:

1. Swapping (interchanging) two rows,
2. Multiplying each element in a row with a number,
3. Adding a multiple of one row to another row.

Result: Transform a system into a equivalent system. Since each row represents an equation, we are essentially **manipulating equations**.

Extension: The application of Gauß-Jordan Algorithm

A system of m equations with n unknowns will have a unique solution if and only if it is *diagonalizable*. i.e. It can be transformed into diagonal form.

Remark: *Diagonalization* turns out to be an important topic in VV286, especially in terms of *ordinary differential equation systems*.

Fundamental Lemma for Homogeneous Equations

The homogeneous system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

of m equations in n real or complex unknowns $x_1, x_2, \ldots, x_n$ has a **non-trivial** solution if $n > m$.

Remark: This fundamental lemma contributes to prove that any basis of a vector space has the same length.

Overview

1. Linear Independence
2. Span
3. Basis
4. Dimension
5. Basis Extension Theorem
6. Sum of Vector Space

Let V be a real or complex vector space and $v_1, v_2, \ldots, v_n \in V$. Then the vectors $v_1, v_2, \ldots, v_n$ are said to be *independent* if for all $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$

$$\sum_{k=1}^{n} \lambda_k v_k = 0 \qquad \Rightarrow \qquad \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0.$$

A finite set $M \subset V$ is called an *independent set* if the elements of M are independent.

Span

Let $v_1, v_2, \ldots, v_n \in V$ and $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$. Then the expression

$$\sum_{k=1}^{n} \lambda_k v_k = \lambda_1 v_1 + \cdots + \lambda_n v_n$$

is called a *linear combination* of the vectors $v_1, v_2, \ldots, v_n$.
The set

$$\operatorname{span}\{v_1, \ldots, v_n\} = \left\{ y \in V : y = \sum_{k=1}^{n} \lambda_k v_k, \lambda_1, \ldots, \lambda_n \in \mathbb{F} \right\}$$

is called the *(linear) span* or the *linear hull* of the vectors $v_1, v_2, \ldots, v_n$.

The vectors $v_1, v_2, \ldots, v_n \in V$ are independent if and only if **none of them is contained in the span of all the others**.

(How to prove?)

Basis

Let V be a real or complex vector space. An n-tuple $\mathcal{B} = (b_1, \dots, b_n) \in V^n$ is called an *(ordered and finite) basis* of V if every vector v has a **unique** representation

$$v = \sum_{i=1}^{n} \lambda_i b_i, \qquad \lambda_i \in \mathbb{F}. \tag{4}$$

The numbers λ_i are called the *coordinates* of v with respect to $\mathcal{B}$.

The tuple of vectors $(e_1, e_2, \dots, e_n)$, $e_i \in \mathbb{R}^n$,

$$e_i = (0, \dots 0, \underset{\substack{\uparrow \\ i\text{th} \\ \text{entry}}}{1}, 0, \dots, 0), \qquad i = 1, \dots, n,$$

is called the *standard basis* or *canonical basis* of $\mathbb{R}^n$.

Basis = Independence + Span

Let V be a real or complex vector space.
An n-tuple $\mathcal{B} = (b_1, \dots, b_n) \in V^n$ is a basis of V if and only if

1. the vectors $b_1, b_2, \dots, b_n$ are linearly independent, i.e., $\mathcal{B}$ is an independent set,
2. $V = \operatorname{span} \mathcal{B}$.

(How to prove?)

Remark: This theorem is more practical than the definition of basis when proving some set is a basis of some vector space. It helps one decompose the proof into two parts: 1. prove linear independence (uniqueness of (4)) 2. prove the span is large enough (existence of (4)).

Dimension

Let V be a real or complex finite-dimensional vector space, $V \neq \{0\}$. **Then any basis of V has the same length** (number of elements).

Remark: This theorem can be proved by contradiction (Use the definition of basis and the fundamental lemma for homogeneous equations). With such a premise, we can then define the *dimension* of vector space.

Let V be a real or complex vector space. Then V is called *finite-dimensional* if either

- $V = 0$ or
- V possesses a finite basis.

If V is not finite-dimensional, we say that it is *infinite-dimensional*.

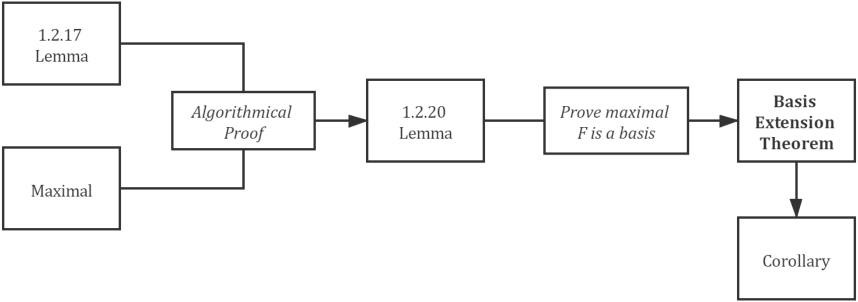

Figure: Logic Flow of Basic Extension Theorem

An interpretation of “**maximal**”: the max (in size) independent subset of some set.

Let V be a finite-dimensional vector space and $A' \subset V$ an independent set. Then there exists a basis of V containing A'.

Remark:
The basis extension theorem is fundamental. It tells us that for any independent subset A' of a finite-dimensional vector space V, we can always find and add $\dim V - |A'|$ elements to A' to extend it into a basis of V. And two useful corollaries follow immediately:

Let V be an n-dimensional vector space, $n \in \mathbb{N}$. Then

1. any independent set A with n elements is a basis of V.
2. an independent set A may have at most n elements.

(How to prove?)

Let V be a real or complex vector space and U, W be sets in V.

(i) We define the *sum of U and W* by

$$U + W := \left\{ v \in V : \underset{u \in U}{\exists} \; \underset{w \in W}{\exists} : v = u + w \right\}.$$

(ii) If U and W are subspaces of V with $U \cap W = \{0\}$, the sum $U + W$ is called *direct*, and we denote it by $U \oplus W$.

Two properties about sum of vector space:

1. The sum $U + W$ of vector spaces U, W is direct if and only if all $x \in U + W$, $x \neq 0$, have a **unique** representation $x = u + w, \; u \in U, w \in W$.
2. Let V be a vector space and $U, W \subset V$ be finite-dimensional subspaces of V. Then

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W.$$

Proof I

Suppose

$$\{v_1, \ldots, v_p\}$$

is a basis for $U \cap W$. By *Basis Extension Theorem*, we can find a basis

$$\{v_1, \ldots, v_p, u_1, \ldots, u_q\}$$

for U and a basis

$$\{v_1, \ldots, v_p, w_1, \ldots, w_r\}$$

for W.
Then we just need to show that

$$B = \{v_1, \ldots, v_p, u_1, \ldots, u_q, w_1, \ldots, w_r\}$$

is a basis for $U + W$

Suppose

$$\alpha_1 v_1 + \cdots + \alpha_p v_p + \beta_1 u_1 + \cdots + \beta_q u_q + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0$$

Then

$$x = \underbrace{\alpha_1 v_1 + \cdots + \alpha_p v_p + \beta_1 u_1 + \cdots + \beta_q u_q}_{\in U} = -(\underbrace{\gamma_1 w_1 + \cdots + \gamma_r w_r}_{\in W})$$

belongs to $U \cap W$. Thus

$$x = \delta_1 v_1 + \cdots + \delta_p v_p$$

and therefore

$$\delta_1 v_1 + \cdots + \delta_p v_p = -(\gamma_1 w_1 + \cdots + \gamma_r w_r)$$

so that

$$\delta_1 v_1 + \cdots + \delta_p v_p + \gamma_1 w_1 + \cdots + \gamma_r w_r = 0$$

Since the set $\{v_1, \ldots, v_p, w_1, \ldots, w_r\}$ is linearly independent, we conclude

$$\delta_1 = 0, \quad \ldots, \quad \delta_p = 0, \quad \gamma_1 = 0, \quad \ldots, \quad \gamma_r = 0$$

and also that

$$\alpha_1 v_1 + \cdots + \alpha_p v_p + \beta_1 u_1 + \cdots + \beta_q u_q = 0$$

So, from linear independence of $\{v_1, \ldots, v_p, u_1, \ldots, u_q\}$ we get

$$\alpha_1 = 0, \quad \ldots, \quad \alpha_p = 0, \quad \beta_1 = 0, \quad \ldots, \quad \beta_q = 0$$

Therefore, the set B is independent. It is clear that $\operatorname{span} B = U + W$. So we conclude B is a basis for $U + W$, and furthermore,

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W.$$

Let V be a vector space and $U, W \subset V$ be finite-dimensional subspaces of V. Then

$$\dim(U + W) \leq \dim U + \dim W.$$

The condition for "$=$": the sum is direct. i.e.

$$\dim(U \oplus W) = \dim U + \dim W.$$

Overview

1. Inner Product Spaces
2. Induced Norm
3. Orthogonality & Orthonormal System
4. The Projection Theorem
5. Gram-Schmidt Orthonormalization

Inner Product Space I

Let V be a real or complex vector space. Then a map $\langle \cdot , \cdot \rangle : V \times V \to \mathbb{F}$ is called a scalar product or inner product if for all $u, v, w \in V$ and all $\lambda \in \mathbb{F}$

1. *Positive-definite*
   $\langle v , v \rangle \geq 0$ and $\langle v , v \rangle = 0$ if and only if $v = 0$,
2. *Linearity in the 2nd argument*
   $\langle u , v + w \rangle = \langle u , v \rangle + \langle u , w \rangle$
3. *Linearity in the 2nd argument*
   $\langle u , \lambda v \rangle = \lambda \langle u , v \rangle$
4. *Conjugate symmetry*
   $\langle u , v \rangle = \overline{\langle v , u \rangle}$

The pair $(V, \langle \cdot , \cdot \rangle)$ is called an *inner product space*.

Prove that

1. $$\langle \lambda u , v \rangle = \overline{\lambda} \langle u , v \rangle .$$

2. $$\langle u + v , w \rangle = \langle u , w \rangle + \langle v , w \rangle$$

This is called the *conjugate linearity* in the 1st argument.

What if $\mathbb{F} = \mathbb{R}$?
Ans: Conjugate symmetry reduces to symmetry, and conjugate linearity reduces to linearity. So, an inner product on a real vector space is a positive-definite symmetric *bilinear map*.

Remark: Multi-linear map will be discussed in detail in *Differential Calculus - Second Derivative*.

Let $(V, \langle \cdot , \cdot \rangle)$ be an inner product space. The map

$$\| \cdot \| : V \to \mathbb{R}, \qquad \|v\| = \sqrt{\langle v , v \rangle}$$

is called the *induced norm* on V.

(How to prove that an induced norm is actually a norm?)

By *the Cauchy-Schwarz inequality*, we define the *angle* $\alpha(u, v) \in [0, \pi]$ *between* u *and* v by

$$\cos \alpha(u, v) = \frac{\langle u , v \rangle}{\|u\| \|v\|}. \tag{5}$$

We are particularly interested in the case that $\alpha = \pi/2$. i.e. $\langle u , v \rangle = 0$. Therefore, we introduce *orthogonality*.

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product vector space.

1. Two vectors $u, v \in V$ are called *orthogonal* or *perpendicular* if $\langle u, v \rangle = 0$. We then write $u \perp v$.
2. We call
$$M^{\perp} := \left\{ v \in V : \underset{m \in M}{\forall} \langle m, v \rangle = 0 \right\}$$
the *orthogonal complement* of a set $M \subset V$.

For short, we sometimes write $v \perp M$ instead of $v \in M^{\perp}$ or $v \perp m$ for all $m \in M$.

Remark: The orthogonal complement $M^{\perp}$ is a subspace of V.
(How to prove?)

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product vector space. A tuple of vectors $(v_1, v_2, \dots, v_r) \in V$ is called a *(finite) orthonormal system* if

$$\langle v_j , v_k \rangle = \delta_{jk} := \begin{cases} 1 & \text{for } j = k, \\ 0 & \text{for } j \neq k, \end{cases} \qquad j, k = 1, \dots, r,$$

i.e., if $\|v_k\| = 1$ and $v_j \perp v_k$ for $j \neq k$.

Let $(V, \langle \cdot, \cdot \rangle)$ be a finite-dimensional inner product vector space and $\mathcal{B} = (e_1, \dots, e_n)$ a basis of V. If $\mathcal{B}$ is also an orthonormal system, we say that $\mathcal{B}$ is an *orthonormal basis* (ONB).

Orthonormal Systems & Bases II

Parseval's Theorem Let $(V, \langle \cdot , \cdot \rangle)$ be a finite-dimensional inner product vector space and $\mathcal{B} = \{e_1, \dots, e_n\}$ an orthonormal basis of V. Then

$$\|v\|^2 = \sum_{i=1}^{n} |\langle v , e_i \rangle|^2$$

for any $v \in V$.

Remark: Parseval's Theorem gives a alternative way to calculate a vector's induced norm.

Let $(V, \langle \cdot , \cdot \rangle)$ be a (possibly infinite-dimensional) inner product vector space and $(e_1, e_2, \ldots, e_r)$, $r \in \mathbb{N}$, be an orthonormal system in V. Denote $U := \operatorname{span}\{e_1, \ldots, e_r\}$.
Then for every $v \in V$ there exists a unique representation

$$v = u + w \qquad\qquad \text{where } u \in U \text{ and } w \in U^{\perp}$$

and $u = \sum_{i=1}^{r} \langle e_i , v \rangle e_i$, $w := v - u$. The vector

$$\pi_U v := \sum_{i=1}^{r} \langle e_i , v \rangle e_i$$

is called the *orthogonal projection* of v onto U.

The projection theorem essentially states that $\pi_U v$ **always exists** and is independent of the choice of the orthonormal system (it **depends only on the span** U of the system).

Moreover, it generalize the idea of projection:

$$\pi_{e_i} v \longrightarrow \pi_U v$$

A vector in an inner product space can be decomposed not only on its orthonormal basis but also on its subspaces.

Orthonormal System $\sim$ Best Approximation

Why is orthonormal system extremely useful?

Let V be a (infinite) vector space. We can approximate an element $v \in V$ using a (finite) linear combination of some orthonormal basis. This is useful in engineering problems.

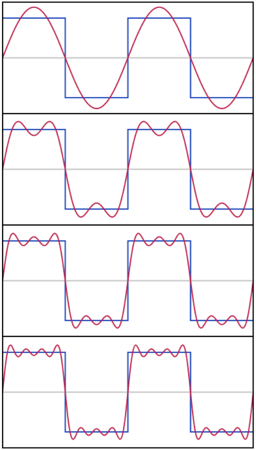

Figure: The Fourier Series Approximation of A Square Wave

Gram-Schmidt Orthonormalization

Just remember how to do it.

How to use Gram-Schmidt Orthonormalization to obtain *Legendre polynomials*?