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 26, 2020

Something you need to pay attention to...

Think More and Be Interactive!

- Do think more about the question in "()".
  e.g. "(How to prove?)"
- You are welcome to ask questions in a adequate manner.
- Please open your camera so that I can receive more feedbacks from you. (Makes our life easier!)
- The class is designed to be interactive. However, if you really do not want to be asked at all, please type an "_" before your zoom name.

Overview

1. Definition of Linear Maps
2. Homomorphism
3. **Range & Kernel**
4. Dual Basis
5. Dimension Formula
6. **Operator Norm**

Definition of Linear Maps

Let $(U, \oplus, \odot)$ and $(V, \boxplus, \boxdot)$ be vector spaces that are either both real or both complex. Then a map $L : U \to V$ is said to be *linear* if it is both *homogeneous*, i.e.,

$$L(\lambda \odot u) = \lambda \boxdot L(u)$$

and *additive*, i.e.,

$$L(u \oplus u') = L(u) \boxplus L(u'),$$

for all $u, u' \in U$ and $\lambda \in \mathbb{F}$. The set of all linear maps $L : U \to V$ is denoted by $\mathcal{L}(U, V)$.

Definition of Linear Maps

Which of them are linear maps?

1. For $I \subset \mathbb{R}$, the map $\frac{d}{dx} : f \mapsto f'$;
2. For $(0, 1)$, the map $T : f \mapsto \int_0^1 f$;
3. If $\mathbb{C}$ is regarded as a real vector space, the map $z \mapsto \overline{z}$;
4. If $\mathbb{C}$ is regarded as a complex vector space, the map $z \mapsto \overline{z}$;
5. For polynomials p, the map $T : p \mapsto (Tp)(x) = x^2 p(x)$;
6. $T : \mathbb{R}^2 \to \mathbb{R}$, $T(x, y) = \sqrt{xy}$;
7. $T : \mathbb{R} \to \mathbb{R}$, $T(x) = x + 1$;
8. A continuous additive map;

Definition of Linear Maps

Which of them are linear maps?

1. For $I \subset \mathbb{R}$, the map $\frac{d}{dx} : f \mapsto f'$;
2. For $(0,1)$, the map $T : f \mapsto \int_0^1 f$;
3. If $\mathbb{C}$ is regarded as a real vector space, the map $z \mapsto \overline{z}$;
4. If $\mathbb{C}$ is regarded as a complex vector space, the map $z \mapsto \overline{z}$;
5. For polynomials p, the map $T : p \mapsto (Tp)(x) = x^2 p(x)$;
6. $T : \mathbb{R}^2 \to \mathbb{R}$, $T(x,y) = \sqrt{xy}$;
7. $T : \mathbb{R} \to \mathbb{R}$, $T(x) = x + 1$;
8. A continuous additive map;

Extension: Continuous and additive implies linear

Homomorphism

According to their properties, there are several fancy names for linear maps. A homomorphism $L \in \mathcal{L}(U, V)$ is said to be

- an isomorphism if L is bijective;
- an endomorphism if $U = V$;
- an automorphism if $U = V$ and L is bijective;
- epimorph if L is surjective;
- monomorph if L is injective.

Remark: They are just fancy names. You only need to know what is an isomorphism.

Homomorphism & Finite-Dimensional Spaces

Let U, V be real or complex vector spaces and $(b_1, b_2, \ldots, b_n)$ a basis of U (in particular, it is assumed that $\dim U = n < \infty$). Then for every n-tuple $(v_1, v_2, \ldots, v_n) \in V^n$ there exists a unique linear map $L : U \to V$ such that $Lb_k = v_k, \; k = 1, \ldots, n$.

Let U, V be finite-dimensional vector spaces and $L \in \mathcal{L}(U, V)$. Then L is an isomorphism if and only if for every basis $(b_1, b_2, \ldots, b_n)$ of U the tuple $(Lb_1, Lb_2, \ldots, Lb_n)$ is a basis of V. [L generates a basis of V from U]

(How to prove?)

Proof

($\Rightarrow$) Assume that L is bijective. Then for $y \in V$ the pre-image $x = L^{-1}y$ is uniquely determined. Let $x = \sum \lambda_k b_k$ be the representation of x in the basis $\mathcal{B} = (b_1, \dots, b_n)$. Now

$$y = L\left(\sum_{k=1}^{n} \lambda_k b_k\right) = \sum_{k=1}^{n} \lambda_k \cdot Lb_k$$

where the λ_k are uniquely determined by x, which is uniquely determined by y. Thus for any y we can find a representation in terms of $(Lb_1, Lb_2, \dots, Lb_n)$ by considering the pre-image $x = L^{-1}y$.
We still need to show that this representation is unique, i.e., if $y = \sum \mu_k \cdot Lb_k$, then $\mu_k = \lambda_k$. Applying L^{-1}, we see that

$$L^{-1}y = x = \sum_{k=1}^{n} \lambda_k b_k, \qquad L^{-1}y = L^{-1}\sum_{k=1}^{n} \mu_k \cdot Lb_k = \sum_{k=1}^{n} \mu_k b_k$$

and because $(b_1, b_2, \dots, b_n)$ is a basis we see that $\mu_k = \lambda_k$.

Question: Is the proof of uniqueness necessary?

Proof

($\Rightarrow$) Assume that L is bijective. Then for $y \in V$ the pre-image $x = L^{-1}y$ is uniquely determined. Let $x = \sum \lambda_k b_k$ be the representation of x in the basis $\mathcal{B} = (b_1, \ldots, b_n)$. Now

$$y = L\left(\sum_{k=1}^{n} \lambda_k b_k\right) = \sum_{k=1}^{n} \lambda_k \cdot Lb_k$$

where the λ_k are uniquely determined by x, which is uniquely determined by y. Thus for any y we can find a representation in terms of $(Lb_1, Lb_2, \ldots, Lb_n)$ by considering the pre-image $x = L^{-1}y$.
We still need to show that this representation is unique, i.e., if $y = \sum \mu_k \cdot Lb_k$, then $\mu_k = \lambda_k$. Applying L^{-1}, we see that

$$L^{-1}y = x = \sum_{k=1}^{n} \lambda_k b_k, \qquad L^{-1}y = L^{-1}\sum_{k=1}^{n} \mu_k \cdot Lb_k = \sum_{k=1}^{n} \mu_k b_k$$

and because $(b_1, b_2, \ldots, b_n)$ is a basis we see that $\mu_k = \lambda_k$.

Question: Is the proof of uniqueness necessary?

We can by some method specify some solution of λ_k uniquely. However, **it does not mean we cannot use another method to find other valid values for λ_k.** Thus we still need to prove the uniqueness of the representation!

Prove Uniqueness and Existence

To prove A is the *unique* element that satisfy condition P:

- Assume A, B both satisfy P, and prove $A = B$.

To prove there *exists* an element A that satisfy condition P:

- Find an explicit representation for A that satisfy P.
- Give an example.
- Provide an algorithm to generate such A.

Dual Basis I

Let V be a real or complex vector space. Then $\mathcal{L}(V, \mathbb{F})$ is known as the *dual space* of V and denoted by V^*. The dual space of V is of course itself a vector space.

Let $\dim V = n < \infty$ and $\mathcal{B} = (b_1, \dots, b_n)$ be a basis of V. Then for every $k = 1, \dots, n$ there exists a unique map

$$b_k^* : V \to \mathbb{F}, \qquad\qquad b_k^*(b_j) = \delta_{jk} = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases}$$

It turns out (see exercises) that the tuple of maps $B^* = (b_1^*, \dots, b_n^*)$ is a basis of $V^* = \mathcal{L}(V, \mathbb{F})$ (called the *dual basis of* $\mathcal{B}$) and thus $\dim V^* = \dim V = n$. (see Assignment 2)

Dual Basis II

Example: The dual basis for $\mathbb{R}^2$ whose basis are $\{e_1, e_2\}$:

$$\{(1 \quad 0), (0 \quad 1)\}$$

Question: Why on earth do we care dual space?

In brief, just as concrete vectors $(x_1, \ldots, x_n)^T \in \mathbb{R}^n$ are naturally generalized to elements of vector spaces, concrete linear expressions $a_1x_1 + \ldots + a_nx_n$ in $x_1, \ldots, x_n$ are naturally generalized to linear functionals.

Extension: Why do we care dual space?

Range and Kernal

Let U, V be real or complex vector spaces and $L \in \mathcal{L}(U, V)$. Then we define the range of L by

$$\operatorname{ran} L := \left\{ v \in V : \underset{u \in U}{\exists}\ v = Lu \right\}$$

and the *kernel* of L by

$$\ker L := \{ u \in U : \ Lu = 0 \} .$$

It is easy to see that $\operatorname{ran} L \subset V$ and $\ker L \subset U$ are subspaces.

Exercise

Prove: $L \in \mathcal{L}(U, V)$ is injective if and only if $\ker L = \{0\}$.

Exercise

Prove: $L \in \mathcal{L}(U, V)$ is injective if and only if $\ker L = \{0\}$.

We prove

$$Lu_1 = Lu_2 \Rightarrow u_1 = u_2$$

$$\Leftrightarrow$$

$$Lu = 0 \Rightarrow u = 0$$

Remark: This relation is useful when proving injective.

Isomorphism

Two finite-dimensional vector spaces U and V are isomorphic if and only if they have the same dimension:

$$U \cong V \qquad \Leftrightarrow \qquad \dim U = \dim V$$

This is a fundamental result, establishing the foundation for *calculus of linear algebra* (matrices theorem).

Dimension Formula

A deep and fundamental result on linear algebra:

Let U, V be real or complex vector spaces, $\dim U < \infty$. Let $L \in \mathcal{L}(U, V)$. Then

$$\dim \operatorname{ran} L + \dim \ker L = \dim U. \tag{1}$$

Corollary

Let U, V be real or complex finite-dimensional vector spaces with $\dim U = \dim V$. Then a linear map $L \in \mathcal{L}(U, V)$ is injective if and only if it is surjective.

(How to prove?)

Remark: We encounter an alternative, either

- L is bijective, or
- L is nor surjective or injective.

Bounded Linear Maps

5.16. Definition. Let $(U, \|\cdot\|_U)$ and $(V, \|\cdot\|_V)$ be normed vector spaces. Then a linear map $L\colon U \to V$ is said to be ***bounded*** if there exists some constant $c > 0$ (called a ***bound*** for L) such that

$$\|Lu\|_V \le c \cdot \|u\|_U \qquad \text{for all } u \in U. \tag{5.6}$$

Remark: It can be shown that if U is a finite-dimensional vector space, then any linear map is bounded. (You will prove this in VV286, following from the fact that **any two norms on a finite-dimensional space are equivalent**.)

Operator Norm

1.4.19. Definition and Theorem. Let U, V be normed vector spaces. Then the set of bounded linear maps $\mathscr{L}(U, V)$ is also a vector space and

$$\|L\| := \sup_{\substack{u \in U \\ u \neq 0}} \frac{\|Lu\|_V}{\|u\|_U} = \sup_{\substack{u \in U \\ \|u\|_U = 1}} \|Lu\|_V. \tag{1.4.7}$$

defines a norm, the so-called ***operator norm*** or ***induced norm*** on $\mathscr{L}(U, V)$.

The proof of the norm properties is left to the reader. The operator norm also has the additional, very useful, property that

$$\|L_2 L_1\| \leq \|L_2\| \cdot \|L_1\|, \qquad L_1 \in \mathscr{L}(U, V), \quad L_2 \in \mathscr{L}(V, W).$$

Remark: Operator norm is quite useful due to its good property. Try to prove this property.

Overview

1. Definition of Matrix
2. **Matrices as Linear Maps**
3. Matrix Product
4. Matrix Transpose/Adjoint
5. Property of Matrices
6. Matrix Inverse
7. Change of Basis

Definition of Matrix

Since that

- Every n-dimensional real vector space is isomorphic to $\mathbb{R}^n$;
- Every n-dimensional complex vector space is isomorphic to $\mathbb{C}^n \cong \mathbb{R}^{2n}$;

we only need to focus on $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \cong \mathrm{Mat}(m \times n; \mathbb{F})$.
An $m \times n$ matrix over the matrix over the complex numbers is a map

$$a : \{1, \dots, m\} \times \{1, \dots, n\} \to \mathbb{C}, \qquad (i, j) \mapsto a_{ij}.$$

We represent the graph of a through

$$A := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}.$$

Matrices = Linear Maps

We consider matrices and linear maps as *actually identical* because

- every linear map between finite-dimensional vector spaces may be expressed as a matrix, and
- every matrix corresponds (in a certain way) to some such linear map.

Each matrix $A \in \mathrm{Mat}(m \times n; \mathbb{R})$ uniquely determines a linear map $j(A) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ such that **the columns $a_{\cdot k}$ are the images of the standard basis vectors $e_k \in \mathbb{R}^n$**; in particular,

$$j : \mathrm{Mat}(m \times n; \mathbb{R}) \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$$

is an isomorphism, $\mathrm{Mat}(m \times n; \mathbb{R}) \cong \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, so every map $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ corresponds to a matrix $j^{-1}(L)$ whose columns $a_{\cdot k}$ are the images of the standard basis vectors $e_k \in \mathbb{R}^n$.

Remark: This theorem is extremely important for us to understand the essence of matrices —— **They are nothing special but just linear maps** defined in certain way. Recall the sentence on the cover page of my slide:

"Matrices are just linear maps!"

Why is this equivalency so important? Because inspired by this theorem we can further define the operation of matrices, prove some cool properties of matrices, and etc. It will greatly help you understand what's going on in the future study of linear algebra.
For example, we will soon show that the derivatives of a vector function is actually a matrix! Why is that? Recall that the derivative is the linear approximation of the original function at some point. And a matrix is just a linear map! Then it immediately makes sense.

Matrices = Linear Maps

Matrix Product

Let $A \in \text{Mat}(l \times m; \mathbb{C})$ and $B \in \text{Mat}(m \times n; \mathbb{C})$. Then the product of A and B is

$$\left(Ab_1, Ab_2, ..., Ab_j\right)$$

The result of matrix multiplication is that we apply A to each column of B.

Matrix Transpose/Adjoint

For $A = (a_{ij}) \in \mathrm{Mat}(m \times n; \mathbb{F})$ we define the *transpose* of A by

$$A^T \in \mathrm{Mat}(n \times m; \mathbb{F}), \qquad A^T = (a_{ji}).$$

We also define the *adjoint*

$$A^* \in \mathrm{Mat}(n \times m; \mathbb{F}), \qquad A^* = \overline{A}^T = (\overline{a_{ji}}).$$

where in addition to the transpose the complex conjugate of each entry is taken.

It is easy to see (in the assignments) that for $A \in \mathrm{Mat}(m \times n; \mathbb{F}),\ x \in \mathbb{F}^m,\ y \in \mathbb{F}^n$,

$$\langle x, Ay \rangle = \langle A^* x, y \rangle.$$

Matrix Properties

- Not commutative for product. $AB \neq BA$
- Associative for product. $(AB)C = A(BC)$
- commutative for sum. $A + B = B + A$
- Associative for sum. $A + B + C = A + (B + C)$
- Right and left distributive for product. $A(B + C) = AB + AC$ and $(D + E)F = DF + EF$
- Transpose of product. $(AB)^T = B^T A^T$
- Transpose of sum. $(A + B)^T = A^T + B^T$
- Transpose & Inverse. $(A^T)^{-1} = (A^{-1})^T$
- Inverse of two invertible matrix's product (if exists) $(AB)^{-1} = B^{-1}A^{-1}$

Matrix Inverse

A matrix $A \in \mathrm{Mat}(n \times n; \mathbb{R})$ is called *invertible* if there exists some $B \in \mathrm{Mat}(n \times n; \mathbb{R})$ such that

$$AB = BA = \mathrm{id} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}. \tag{2}$$

We then write $B = A^{-1}$ and say that A^{-1} is the *inverse* of A.

Remark:

- The inverse is unique. (How to prove?)
- We only need to verify $AB = id$ or $BA = id$ to conclude $B = A^{-1}$ by *6.12. Remark.*
- However, it is not always true (For finite-dimensional vector space, it is true) that $L_1 L_2 = id \Rightarrow L_2 L_1 = id$ for two general linear maps L_1, L_2!

Left Inverse & Right Inverse

Let l^2 be space of square summable sequence and L, R be left and right shift defined by

$$L : l^2 \to l^2, (a_n) \mapsto (l_n) = (a_{n+1})$$
$$R : l^2 \to l^2, (a_n) \mapsto (r_n) = \begin{cases} 0 & n = 0 \\ a_{n-1} & n \neq 0 \end{cases}$$

We notice that $RL = id \neq LR$. Hence we say that for L, only *left inverse* exists; for R, only *right inverse* exists.

Question: What is the adjoint of L? (Hint: use the definition of adjoint in your assignment.)

Change of Basis