VV285 RC Part II

Elements of Linear Algebra

"Linear Algebra!"

Yiwen Tu

Univerity of Michigan-Shanghai Jiao Tong University Joint Institute

May 25, 2022

JOINT INSTITUTE
交大密西根学院

Matrices

5.1. Definition. An $m \times n$ matrix over the complex numbers is a map

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

We represent the graph of a through

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

We denote the set of all $m \times n$ matrices over $\mathbb{C}$ by $\mathrm{Mat}(m \times n; \mathbb{C})$.

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

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

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

Matrices

5.4. Definition. Let $A \in \mathrm{Mat}(l \times m; \mathbb{C})$ and $B \in \mathrm{Mat}(m \times n; \mathbb{C})$. Then we define the ***product of $A = (a_{ik})$ and $B = (b_{kj})$*** by

$$AB \in \mathrm{Mat}(l \times n; \mathbb{C}), \qquad AB := \Big(\sum_{k=1}^{m} a_{ik} b_{kj}\Big)_{\substack{i=1,\dots,l \\ j=1,\dots,n}}$$

We have seen that the matrix product satisfies $j(A) \circ j(B) = j(AB)$. Furthermore, the product is ***associative***, i.e.,

$$\begin{aligned} A(BC) &= j^{-1}(j(A) \circ j(BC)) = j^{-1}(j(A) \circ (j(B) \circ j(C))) \\ &= j^{-1}((j(A) \circ j(B)) \circ j(C)) = j^{-1}(j(AB) \circ j(C)) \\ &= (AB)C \end{aligned}$$

If $A, B \in \mathrm{Mat}(n \times n; \mathbb{C})$ both products AB and BA exist; however

$$AB \neq BA,$$

so the matrix product is ***not commutative***.

Matrices

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

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

For example,

$$\begin{pmatrix} 5 & 6 & 7 \\ 1 & 0 & 2 \end{pmatrix}^T = \begin{pmatrix} 5 & 1 \\ 6 & 0 \\ 7 & 2 \end{pmatrix}.$$

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 $A \in \mathrm{Mat}(m \times n; \mathbb{F})$, $x \in \mathbb{F}^m$, $y \in \mathbb{F}^n$,

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

5.8. Elementary Matrix Manipulations. An elementary row manipulation of a matrix is one of the following:

(i) Swapping (interchanging) of two rows,

(ii) Multiplication of a row with a non-zero number,

(iii) Addition of a multiple of one row to another row.

We are now able to properly define the matrix of a linear map between two finite-dimensional vector spaces.

Let U, V be finite-dimensional real or complex vector spaces with bases

$$\mathscr{A} = (a_1, \dots, a_n) \subset U \qquad \text{and} \qquad \mathscr{B} = (b_1, \dots, b_m) \subset V.$$

Define the isomorphisms

$$\begin{aligned} \varphi_{\mathscr{A}} &: U \xrightarrow{\cong} \mathbb{R}^n, & \varphi_{\mathscr{A}}(a_j) &= e_j, & j &= 1, \dots, n, \\ \varphi_{\mathscr{B}} &: V \xrightarrow{\cong} \mathbb{R}^m, & \varphi_{\mathscr{B}}(b_j) &= e_j, & j &= 1, \dots, m. \end{aligned}$$

Then any linear map $L \in \mathscr{L}(U, V)$ induces a matrix $A = \Phi^{\mathscr{B}}_{\mathscr{A}}(L)$ $\in \text{Mat}(m \times n; \mathbb{R})$ through

$$\begin{tikzcd} U \arrow[r, "L"] \arrow[d, "\varphi_{\mathscr{A}}"'] & V \arrow[d, "\varphi_{\mathscr{B}}"] \\ \mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^m \end{tikzcd} \qquad \Phi^{\mathscr{B}}_{\mathscr{A}}(L) = A = \varphi_{\mathscr{B}} \circ L \circ \varphi_{\mathscr{A}}^{-1}$$

5.9. Definition. 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

$$AB = BA = \mathrm{id} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}. \tag{5.3}$$

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

5.13. Lemma. Let $A \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^n)$. Then A is ***invertible*** if and only if there exists an elementary matrix S corresponding to elementary row operations that transform A into the unit matrix $SA = \mathrm{id}$.

$$
\begin{array}{c|cl}
SA & S & \\
\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{array}{l} :2 \quad \cdot(-2) \\ \quad + \end{array} \\
\begin{pmatrix} 1 & 3/2 \\ 0 & -2 \end{pmatrix} & \begin{pmatrix} 1/2 & 0 \\ -1 & 1 \end{pmatrix} & \begin{array}{l} \quad + \\ :(-2) \quad \cdot(-3/2) \end{array} \\
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \underbrace{\begin{pmatrix} -1/4 & 3/4 \\ 1/2 & -1/2 \end{pmatrix}}_{=A^{-1}} &
\end{array}
$$

5.15. Remark. We note that if $A, B \in \mathrm{Mat}(n \times n; \mathbb{R})$ are invertible, then so is their product $AB \in \mathrm{Mat}(n \times n; \mathbb{R})$ and $(AB)^{-1} = B^{-1}A^{-1}$.

We can use this procedure to find the inverse of any vector space isomorphism L:

$$\begin{tikzcd} U \arrow[r, "L"] \arrow[d, "\varphi_{\mathscr{A}}"'] & V \arrow[d, "\varphi_{\mathscr{B}}"] \\ \mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^m \end{tikzcd}$$

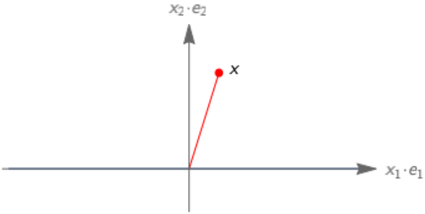
$$L^{-1} = \varphi_{\mathscr{A}}^{-1} \circ A^{-1} \circ \varphi_{\mathscr{B}}$$

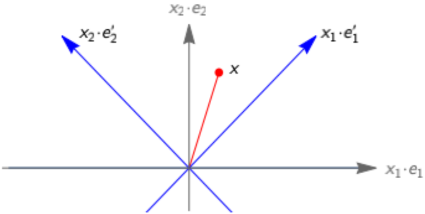
5.17. Example. Consider a rotation by 45° in the clockwise direction,

$$T : \mathbb{R}^2 \to \mathbb{R}^2, \qquad \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

so

$$T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$





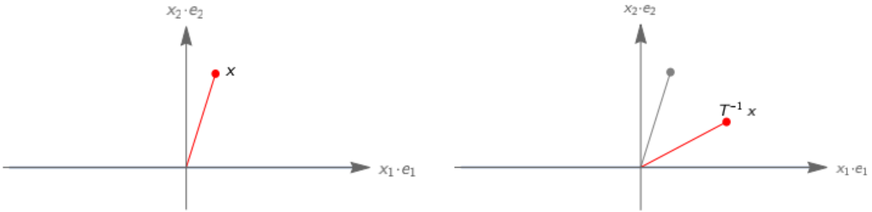
Suppose that

$$x = \sum_{i=1}^{n} x_i e_i = \sum_{i=1}^{n} x_i' e_i' = \sum_{i=1}^{n} x_i' T e_i.$$

Then

$$T^{-1}x = \sum_{i=1}^{n} x_i' e_i$$

and we can find the coordinates x_i', $i = 1, \ldots, n$, of x with respect to $\mathscr{B}'$ simply by applying T^{-1}.



Invertibility of matrices

Suppose $||A - I|| < 1$. Show that A is invertible. In particular, show that if

$$A - \alpha I < ||A||, \alpha \geq ||A||,$$

then A is invertible.

Invertibility of matrices

This exercise describes the “size” of the set of non-invertible operators. Suppose that A is invertible. Show that there is some $\epsilon > 0$ such that for any $||A - B|| < \epsilon$, B is invertible.

Operator Norm

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

$$\|L\| := \sup_{\substack{u \in U \\ u \neq 0}} \frac{\|Lu\|_V}{\|u\|_U} = \sup_{\substack{u \in U \\ \|u\|_U = 1}} \|Lu\|_V. \tag{4.7}$$

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

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

Adjoint

Adjoint Suppose $T \in L(V, W)$. The adjoint of T is the function $T^* : W \to V$ such that

$$\langle Tv, w\rangle = \langle v, T^*\rangle$$

for every $v \in V$, $w \in W$

The definition makes sense, suppose $T \in L(V, W)$. fix $w \in W$. Consider the linear functional on V that maps $v \in V$ to $\langle Tv, w\rangle$; this linear functional depends on T and w. By the Riesz representation Theorem, there exists a unique vector in V such that this linear functional is given by takeing the inner product with it. We call this unique vector T^*w. In other words, T^*w is the unique vector in V such that the requirement is met for every v in V.

Properties of the adjoint

- $(S+T)^* = S^* + T^*$ for all $S, T \in L(V, W)$
- $(\lambda T)^* = \bar{\lambda} T^*$ for all $\lambda \in F$ and $T \in L(V, W)$
- $(T^*)^* = T$ for all $T \in L(V, W)$
- $I^* = I$, where I is the identity operator on V;
- $(ST)^* = T^* S^*$ for all $T \in L(V, W)$ and $S \in L(W, U)$ (here U is an inner product space over F)

Kernel and Range

Suppose $T \in L(V, W)$. then

- $ker T^* \perp ran T$
- $ker T \perp ran T^*$

Conjugate Transpose

The *conjugate transpose* of an m-by-n matrix is the n-by-m matrix obtained by interchanging the rows and columns and then taking tha complex conjugate of each entry. Let $T \in L(V, W)$. Suppose $e_1, \cdots e_n$ is an orthonormal basis of V and $f_1, \cdots, f_m$ is an orthonormal basis of W. then

$$M(T^*, (f_1, \cdots, f_m), (e_1, \cdots, e_n))$$

is the conjugate transpose of

$$M(T, (e_1, \cdots, e_n), (f_1, \cdots, f_m))$$

The proof is not important.

Self-adjoint operators

An operator $T \in L(V)$ is called *self-adjoint* if $T = T^*$. In other words, $T \in L(V)$ is self-adjoint if and only if

$$\langle Tv, w \rangle = \langle v, Tw \rangle$$

for all $v, w \in V$ You'll run into more details of self-adjoint operators in VV286!

Try to prove the following two statements!

- Suppose V is a complex inner product space and $T\ inL(V)$. Suppose
$$\langle Tv, v\rangle = 0$$
for all $v \in V$. Then T=0
- Suppose V is a complex inner product space and $T\ inL(V)$. Then T is self-adjoint if and only if
$$\langle Tv, v\rangle \in R$$
for all $v \in V$.

Normal operators

An operator on an inner product space is called *normal* if it commutes with its adjoint, or:

$$TT^* = T^*T$$

Obviously every self-adjoint operator is normal.

An operator $T \in L(V)$ is normal if and only if

$$||Tv|| = ||T^*v||$$

Learn Well

And

Have Fun!