Linear Algebra I

Chapter 4. Linear Transformations

Boqing Xue

WELCOME!

2023 Fall

上海科技大学 数学科学研究所
Institute of Mathematical Sciences, ShanghaiTech University

Definition of Linear Transformation

Recall: For a matrix $A \in M_{m\times n}(\mathbb{R})$, the matrix transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is given by

$$T_A(\mathbf{x}) = A\mathbf{x}, \quad (\forall \mathbf{x} \in \mathbb{R}^n).$$

It satisfies that

$$T_A(\mathbf{u}+\mathbf{v}) = T_A(\mathbf{u}) + T_A(\mathbf{v}), \quad T_A(k\mathbf{u}) = kT_A(\mathbf{u})$$

for any $\mathbf{u}, \mathbf{u} \in \mathbb{R}^n$ and any scalar k.

Definition. Let V, W be $\mathbb{F}$-vector spaces. A map $T : V \to W$ is called a linear transformation if the following two properties hold for all vectors $\mathbf{u}, \mathbf{v} \in V$ and for all scalars $k \in \mathbb{F}$:

(i) $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$

(ii) $T(k\mathbf{u}) = kT(\mathbf{u})$.

Remark: ◇ When $V = W$, we also call T a linear operator on V.

◇ (i)+(ii) $\iff T(k\mathbf{u}+\ell\mathbf{v}) = kT(\mathbf{u}) + \ell T(\mathbf{v})$ ($\forall \mathbf{u}, \mathbf{v} \in V$ and $k, \ell \in \mathbb{F}$).

◇ $T(\mathbf{0}) = \mathbf{0}$.

◇ $T(k_1\mathbf{u}_1 + \dots k_n\mathbf{u}_n) = k_1T(\mathbf{u}_1) + \dots + k_nT(\mathbf{u}_n)$.

Examples of Linear Transformations

Example. The matrix Transformations $T_A : \mathbb{C}^n \to \mathbb{C}^m$, $\mathbf{x} \mapsto A\mathbf{x}$ with $A \in M_{m\times n}(\mathbb{C})$ are linear.

Example. For any vector spaces V and W, the zero transformation is defined by $0: V \to W$ is defined by

$$0(\mathbf{u}) = \mathbf{0}, \quad (\forall \mathbf{u} \in \mathbf{V}).$$

Example. Let V be a real vector space and $k \in \mathbb{R}$. Then $T: V \to V$ defined by

$$T(\mathbf{u}) = k\mathbf{u}, \quad (\forall \mathbf{u} \in V)$$

is a linear operator on V.

When $k = 0$, it is the zero operator.
When $k = 1$, it is called the identity operator and denoted by $T := I$.
When $0 < k < 1$, it is called contraction of V with factor k.
When $k > 1$, it is called dilation of V with factor k.

Examples of Linear Transformations

Example. Recall that P_n is the space of polynomials with real coefficients and of degree $\leq n$. The following two transformations are linear.

- $\diamond$ $M: P_n \to P_{n+1}$ given by $M(p(x)) = xp(x)$;
- $\diamond$ $D: P_{n+1} \to P_n$ given by $D(p(x)) = p'(x)$.

Example. Recall that $C(a, b)$ is the space of real-valued continuous functions on (a, b), and $C^\infty(a, b)$ is the space of real-valued functions on (a, b) with derivatives of all orders. The following two operators are linear.

- $\diamond$ $D: C^\infty(a, b) \to C^\infty(a, b)$ given by $D(f(x)) = f'(x)$;
- $\diamond$ $S: C(a, b) \to C(a, b)$ given by

$$S(f(x)) = \int_a^x f(t)dt, \quad (a < x < b).$$

Examples of Linear Transformations

Example. Let V be the space of all real Cauchy sequences. The transformation $\mathrm{Lim}: V \to \mathbb{R}$ defined by

$$\mathrm{Lim}(\{a_n\}) = \lim_{n\to\infty} a_n, \quad (\forall \{a_n\} \in V)$$

is linear.

Examples of Linear Transformations

Example. The map $T_1 : \mathbb{R}^3 \to \mathbb{R}^3$ and $T_2 : \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$T_1(x, y, z) = (x, y, 0), \quad T_2(x, y, z) = (x, y)$$

are linear transformations. The effect of these transformations are projection on the xOy-plane in $\mathbb{R}^3$.

Now we step further to infinite-dimensional spaces (countable case).

Example. Recall that

$$\mathbb{R}^{\mathbb{N}} = \{(x_1, x_2, \ldots, x_n, x_{n+1}, \ldots) : x_i \in \mathbb{R}\,(i = 1, 2, \ldots)\}$$

is the space of all sequences. Given indices $i_1, i_2, \ldots, i_r \in \mathbb{N}$, the projection $T : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^r$ on the given components, given by

$$T(x_1, x_2, \ldots, x_n, \ldots) = (x_{i_1}, x_{i_2}, \ldots, x_{i_r}),$$

is a linear transformation.

For example, taking $r = 3$ and $i_1 = 1, i_2 = 2, i_3 = 3$, we have

$$T(1, 2, 3, 4, \ldots, n, \ldots) = (1, 2, 3),$$
$$T(1, -1, 1, -1, \ldots, (-1)^n, \ldots) = (1, -1, 1).$$

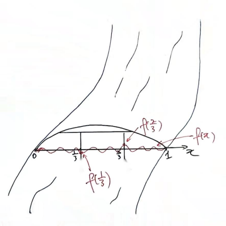
Examples of Linear Transformations

Next we step further to infinite-dimensional spaces (uncountable case).

Example. Let V be any subspace of $F(a,b)$, the space of all functions on (a,b). Take $x_1, x_2, \ldots, x_r \in (a,b)$. The evaluation transformation on V at $x_1, x_2, \ldots, x_r$ is linear. Indeed, it is given by $T : V \to \mathbb{R}^r$, where

$$T(f) = (f(x_1), f(x_2), \ldots, f(x_r)) \quad (\forall f \in V).$$

A bridge on a river (one side at 0 and the other side at 1) has two piers (at $x_1 = 1/3$ and $x_2 = 2/3$). The height of the surface of the river can be viewed as a continuous function on $(0,1)$. To obtain the height at the piers, we take $r = 2$ and above x_1, x_2. For example



$$T(0.2\cos(99\pi x)) = (0.2\cos(33\pi), 0.2\cos(66\pi)) = (-0.2, 0.2).$$

Examples of Linear Transformations

Example. The followings give For a given vector $\mathbf{v}_0 \in \mathbb{R}^n$, we have the linear transformation $T : \mathbb{R}^n \to \mathbb{R}$ given by

$$T(\mathbf{u}) = \mathbf{u} \cdot \mathbf{v}_0, \quad (\forall \mathbf{u} \in \mathbb{R}^n).$$

Example. The following two transformations are linear.

◇ $T : M_{m\times n} \to M_{n\times m}$, $A \mapsto A^T$.

◇ $T : M_n(\mathbb{F}) \to \mathbb{F}$, $A \mapsto \mathrm{tr}(A)$.

Example. Given $B, C \in M_n$, we have the followings linear operators on M_n.

◇ $T_{B,C}(A) = BAC$.

◇ $ad_B(A) = BA - AB$.

Remark: We can write $ad_B = T_{B,I} - T_{I,B}$.

Example. The follow transformations are NOT linear.

◇ $T : M_n(\mathbb{R}) \to \mathbb{R}$, $A \mapsto \det(A)$. Here $n \geq 2$.

◇ $T : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{x} \mapsto \mathbf{x} + \mathbf{x}_0$. Here $\mathbf{x}_0 \neq \mathbf{0}$.

Operations on Linear Transformations

Definition. Let $T, T_1, T_2 : V \to W$ be linear transformation, and let c be a scalar. We define the addition and scalar multiplication by

$$(T_1 + T_2)(\mathbf{v}) = T_1(\mathbf{v}) + T_2(\mathbf{v}), \quad (cT)(\mathbf{v}) = c\,T(\mathbf{v}), \quad (\mathbf{v} \in V).$$

Remark: $T_1 + T_2$ and cT above are also linear transformation. Prove it!

Theorem. Let V, W be linear spaces. Then set $L(V, W)$ of all linear transformations from V to W, with the addition and scalar multiplication defined above, is a also a linear space.

Example. Note that $L(\mathbb{R}^n, \mathbb{R}^m) = M_{m\times n}(\mathbb{R})$.

Describing a Linear Transformation

Question: How to describe a linear transform in an efficient way?

★ Philosophy: Whenever we know the information about basis, we can make clear the information of whole space!

- $T : V \to W$: a linear transformation.
- $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$: a basis for V.
- $\widetilde{B} = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$: a basis for W.

- For any $\mathbf{x} \in V$, there are scalars $x_1, x_2, \dots, x_n$ such that

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \dots + x_n\mathbf{v}_n = \sum_{j=1}^{n} x_j\mathbf{v}_j.$$

- Now

$$T(\mathbf{x}) = x_1 T(\mathbf{v}_1) + x_2 T(\mathbf{v}_2) + \dots + x_n T(\mathbf{v}_n) = \sum_{j=1}^{n} x_j T(\mathbf{v}_j).$$

- Since $T(\mathbf{v}_j) \in W$, there are scalars a_{ij} $(1 \le i \le m,\ 1 \le j \le n)$ such that

$$T(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + a_{2j}\mathbf{w}_2 + \dots + a_{mj}\mathbf{w}_m = \sum_{i=1} a_{ij}\mathbf{w}_i$$

for all $1 \le j \le n$.

- Finally,

$$T(\mathbf{x}) = \sum_{i=1}^{m} \sum_{j=1}^{n} \mathbf{w}_i a_{ij} x_j.$$

Describing a Linear Transformation

- Finally,

$$T(\mathbf{x}) = \sum_{i=1}^{m} \sum_{j=1}^{n} \mathbf{w}_i a_{ij} x_j.$$

- Rewirte (Only notation, not real matrices):

$$\mathbf{x} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

$$T(\mathbf{x}) = \begin{bmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \dots & \mathbf{w}_m \end{bmatrix} \left[\begin{array}{c:c:c:c} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

- Two representations of $T(\mathbf{v})$:

$$\begin{aligned} T(\mathbf{x}) &= \begin{bmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \dots & \mathbf{w}_m \end{bmatrix} \left[\begin{array}{c:c:c:c} [T(\mathbf{v}_1)]_{\widetilde{B}} & [T(\mathbf{v}_2)]_{\widetilde{B}} & \dots & [T(\mathbf{v}_n)]_{\widetilde{B}} \end{array}\right] [\mathbf{x}]_B \\ &= \begin{bmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \dots & \mathbf{w}_m \end{bmatrix} [T(\mathbf{x})]_{\widetilde{B}} . \end{aligned}$$

Matrix of a Linear Transformations

Definition. Let $T : V \to W$ be a linear transform. Let $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $\widetilde{B} = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ be bases for V and W, respectively. Then the matrix for T relative to B and $\widetilde{B}$ is defined to be

$$[T]_{\widetilde{B},B} = \left[\ [T(\mathbf{v}_1)]_{\widetilde{B}} \ \vdots\ [T(\mathbf{v}_2)]_{\widetilde{B}} \ \vdots\ \cdots \ \vdots\ [T(\mathbf{v}_n)]_{\widetilde{B}}\ \right].$$

Remark: Now $T : V \to W$ can be determined by

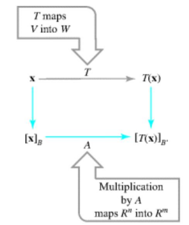
$$[T(\mathbf{x})]_{\widetilde{B}} = [T]_{\widetilde{B},B}\, [\mathbf{x}]_B.$$

Remark: The subscripts $[]_{\widetilde{B},B}$ correspond to $[]_{m\times n}$.

Remark: When $W = V$, we usually choose $\widetilde{B} = B$. Then we write

$$[T]_B = \left[\ [T(\mathbf{v}_1)]_B \ \vdots\ [T(\mathbf{v}_2)]_B \ \vdots\ \cdots \ \vdots\ [T(\mathbf{v}_n)]_B\ \right].$$

Remark: When it makes no confusion, we abbreviate it as $[T]$.



Example. Let $A \in M_{m\times n}(\mathbb{R})$. Let S and $\widetilde{S}$ be the standard bases for $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Then $[T_A]_{\widetilde{S},S} = A$.

Some Examples

Example. (1) Find the matrix for zero transformation $0: V \to W$ relative to any bases.

(2) Find the matrix for the identity operator I on V relative to any basis.

Example. (1) Find the matrix for $T: P_1 \to P_2$ defined by

$$T(p(x)) = xp(3x-5)$$

relative to the standard bases $B = \{1, x\}$ and $\widetilde{B} = \{1, x, x^2\}$.

(2) Compute $T(1+2x)$ (a) directly; (b) by using this matrix.

Solution:

1

Some Examples

Example. Let V be the subspace of $F(-\infty, +\infty)$ spanned by

$$\{e^{2x}, xe^{2x}, x^2e^{2x}, x^3e^{2x}\}.$$

The differentiation operator $D = \frac{d}{dx}$ is a linear operator on V. Find the matrix of D relative to the given basis.

Solution:

2

Some Examples

Example. Let $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. For $ad : M_2 \to M_2$ defined by

$$ad(A) = BA - AB, \quad (A \in M_2),$$

Find the matrix of ad relative to the standard basis for M_2.

Solution:

3

Kernel and Range

Definition. For a linear transformation $T: V \to W$, the kernel and the range (or image) of T are defined by

$$\mathrm{Ker}(T) = \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}\},$$

$$\mathrm{Ran}(T) = \{T(\mathbf{v}) : \mathbf{v} \in V\}.$$

Example. Let $A \in M_{m\times n}(\mathbb{F})$. For the matrix transformation $T_A : \mathbb{F}^n \to \mathbb{F}^m$, we have

$$\mathrm{Ker}(T_A) = \mathrm{Null}(A), \quad \mathrm{Ran}(T_A) = Col(A).$$

Theorem. Let $T: V \to W$ be a linear transformation. Then
- ◇ $\mathrm{Ker}(T)$ is a subspace of V;
- ◇ $\mathrm{Ran}(T)$ is a subspace of W.

Proof:

4

Example. Find the kernel and range of the following linear transformations. Also find the dimensions of theses spaces.

- ◇ The zero operator 0 on a vector space V.
- ◇ The identity operator I on a vector space V.
- ◇ $T : M_n(\mathbb{R}) \to \mathbb{R}$, $A \mapsto \mathrm{tr}(A)$. Here $n \geq 2$.
- ◇ $M : P_n \to P_{n+1}$ given by $M(p(x)) = xp(x)$;
- ◇ $D : P_{n+1} \to P_n$ given by $D(p(x)) = p'(x)$.

Solution:

5

Definition. Let $T : V \to W$ be a linear transform. The rank and nullity of T is defined by

$$\mathrm{rank}(T) = \dim \mathrm{Ran}(T), \quad \mathrm{nullity}(T) = \dim \mathrm{Ker}(T),$$

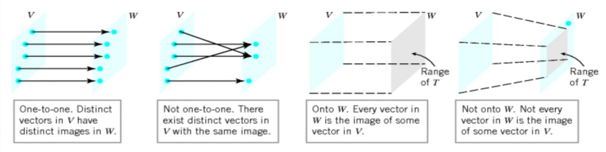
whenever the above subspaces are finite-dimensional.

Theorem. For any linear transformation $T : V \to W$ with V finite-dimensional, we have $\mathrm{rank}(T) + \mathrm{nullity}(T) = \dim(V)$.

Recall: Injective and Surjective maps

Definition. Let $f : V \to W$ be a map from a set V to a set W.

- ⋄ We say that f is injective (one-to-one),
  if it maps distinct elements in V into distinct elements in W.
- ⋄ We say that f is surjective (onto),
  if every element in V is in the range of f.
- ⋄ We say that f is bijective, if it is both injective and surjective.



Definition. A bijective linear transformation from V to W is called an isomorphism. In such case, we say that the two vector spaces V and W are isomorphic.

Remark: They have "same" linear properties!

Examples

Example. One-to-one/Onto/Isomophic or not?

- The zero operator 0 on a vector space V.
- The identity operator I on a vector space V.
- $T : M_n(\mathbb{R}) \to \mathbb{R}$, $A \mapsto \operatorname{tr}(A)$. Here $n \geq 2$.
- $M : P_n \to P_{n+1}$ given by $M(p(x)) = xp(x)$;
- $D : P_{n+1} \to P_n$ given by $D(p(x)) = p'(x)$.
- The right shift T_1 and the left shift T_2 on $\mathbb{R}^{\mathbb{N}}$ given by

$$T_1(x_1, x_2, \ldots, x_n, \ldots) = (0, x_1, \ldots, x_{n-1}, \ldots),$$
$$T_2(x_1, x_2, \ldots, x_n, \ldots) = (x_2, x_3, \ldots, x_{n+1}, \ldots).$$

Solution:

6

Recall: An Equivalence Theorem

Theorem. Suppose that V and W are vector spaces with $\dim(V) = n$, $\dim(W) = m$. Let $T : V \to W$ be a linear transformation.

The following three statements are equivalent:
(1) T is injective;
(2) $\mathrm{Ker}(T) = \{\mathbf{0}\}$; or $\mathrm{nullity}(T) = 0$;
(3) $\mathrm{rank}(T) = n$.

The following three statements are equivalent:
(1) T is surjective;
(2) $\mathrm{Ran}(T) = W$; or $\mathrm{rank}(T) = m$;
(3) $\mathrm{nullity}(T) = n - m$.

When $m = n$, the following three statements are equivalent:
(1) T is bijective (isomorphic);
(2) T is surjective (onto);
(3) T is injective (one-to one).

Remark: If $n > m$, then T cannot be injective.
If $n < m$, then T cannot be surjective.

Remark: The statements in red are NOT correct for INFINITE-dimensional spaces! (See examples in the previous page.)

Theorem. Every n-dimensional $\mathbb{F}$-vector space V is isomorphic to $\mathbb{F}^n$.

Proof:

7

Remark: For any given basis for V, the coordinate map gives an isomorphism.

Example. (1) Find a "natural" isomorphism from P_{n-1} to $\mathbb{R}^n$.
(2) Find a "natural" isomorphism from $M_2(\mathbb{R})$ to $\mathbb{R}^4$.

Solution:

8

Composition of Transformations

Definition. Let U, V, W be three sets and f, g be maps such that

$$U \xrightarrow{f} V \xrightarrow{g} W.$$

Then the composition of g with f, denoted by $g \circ f$, is the map $g \circ f : U \to W$ defined by

$$(g \circ f)(\mathbf{u}) = g(f(\mathbf{u})), \quad (\forall \mathbf{u} \in U).$$

Remark: Indeed, we may also define the multiplication of T_2 and T_1 by $T_2T_1 = T_2 \circ T_1$, and write $T^n = T \circ \ldots \circ T$ for n copies of T.

Theorem. If $T_1 : U \to V$ and $T_2 : V \to W$ are both linear transformations, then $T_2 \circ T_1 : U \to W$ is also a linear transformation.
Proof:

9

Theorem. Let $T_1 : U \to V$ and $T_2 : V \to W$ be linear transformations. Let B, B', B'' be bases for U, V, W, respectively. Then

$$[T_2 \circ T_1]_{B'',B} = [T_2]_{B'',B'}[T_1]_{B',B}.$$

Examples

Example. Find the compositions of linear transformations involving 0 or I.

Example. Find the compositions of T_2 with T_1, where

$$T_1 : P_1 \to P_2,\ p(x) \mapsto xp(x),$$

$$T_2 : P_2 \to P_2,\ p(x) \mapsto p(3x+1).$$

Find $T_2 \circ T_1$ and its matrix relative to the standard bases.

Solution:

10

Examples

Example. Let U be a certain function space on $\mathbb{R}$. Prove the following quantum version of uncertainty principle:

$$P \circ Q - Q \circ P = I.$$

Here P, Q are operators on U defined by

$$(Pf) = f'(x), \quad (Qf)(x) = xf(x), \quad (\forall f \in U).$$

Proof:

11

Remark: P stands for momentum and Q for position.

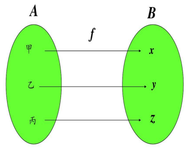
Problem*: Show that there are no operators P and Q on a finite-dimensional space V such that $P \circ Q - Q \circ P = I$.

Inverse of a Transformations

Definition. Let U, V be two sets. Denote by I_U and I_V the identity maps on U and V, respectively. For a map $f : U \to V$, if there is a map $g : V \to U$ such that

$$g \circ f = I_U, \quad f \circ g = I_V,$$

then we say that f is invertible, and call g the inverse of f. Sometimes, we denote the inverse of f by f^{-1}.



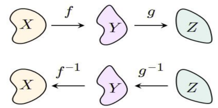
Problem*: Try to prove that a map $f : U \to V$ is bijective, if and only if f is invertible.

Inverse of a Transformations

Theorem. $\diamond$ Let $T : V \to W$ be an invertible linear transformation. Then $T^{-1} : W \to V$ is also a linear transformation.
$\diamond$ Moreover, let B and B' be bases for V and W, respectively. Then $[T]_{B',B}$ is an invertible matrix, and

$$[T^{-1}]_{B,B'} = [T]_{B',B}^{-1}.$$

Attention: In our textbook, the inverse of a one-to-one linear transformation $T : V \to W$ is defined to be $T^{-1} : \mathrm{Ran}(T) \to V$. We do NOT use this definition!



Theorem. Let $T_1 : U \to V$ and $T_2 : V \to W$ be invertible linear transformations. Then $T_2 \circ T_1$ is also invertible, and

$$(T_2 \circ T_1)^{-1} = T_1^{-1} \circ T_2^{-1}.$$

Examples

Example. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear operator defined by

$$T(x_1, x_2, x_3) = (3x_1 + x_2,\ -2x_1 - 4x_2 + 3x_3,\ 5x_1 + 4x_2 - 2x_3).$$

Determine whether T is invertible. If so, find T^{-1}.

Solution:

12

Examples

Example. Let $T : P_2 \to P_2$ be defined by

$$T(p(x)) = p(2x+1).$$

Determine whether T is invertible. If so, find T^{-1}.

Solution:

13

Properties of an Inverse

The following two theorems have already appeared in the version of matrices.

Theorem. Let V and W be two finite-dimensional vector spaces with same dimension. Let $T : V \to W$ be a linear transformation. (See chapter 4.)

- $\diamond$ If T_1 is injective, then T_1 is invertible.
- $\diamond$ If T_1 is surjective, then T_1 is invertible.

Theorem. Let V and W be two finite-dimensional vector spaces with same dimension. Let $T_1 : V \to W$ and $T_2 : W \to V$ be two linear transformations. (See chapter 1.)

- $\diamond$ If $T_2 \circ T_1 = I_V$, then T_1 is invertible, and $T_1^{-1} = T_2$.
- $\diamond$ If $T_2 \circ T_1 = I_V$, then T_2 is invertible, and $T_2^{-1} = T_1$.

Problem*: When V and W are infinite-dimensional, try to find counterexamples of the statements in above two theorems.

Problem*: If $T_2 \circ T_1$ is injective, then T_1 is injective.
If $T_2 \circ T_1$ is surjective, then T_2 is surjective.
Find an example such that $T_2 \circ T_1$ is injective but T_2 is not.
Find an example such that $T_2 \circ T_1$ is surjective but T_1 is not.

Recalling

Definition. Let $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $B' = \{\mathbf{v}'_1, \dots, \mathbf{v}'_n\}$ be two bases of V. The transition matrix from B' to B is defined as

$$P_{B \leftarrow B'} = \left[\begin{array}{c:c:c:c} [\mathbf{v}'_1]_B & [\mathbf{v}'_2]_B & \cdots & [\mathbf{v}'_n]_B \end{array} \right]$$

Proposition. Let $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $B' = \{\mathbf{v}'_1, \dots, \mathbf{v}'_n\}$ be two bases of V. Then for any vector $\mathbf{x} \in V$, it satisfies that

$$[\mathbf{x}]_B = P_{B \leftarrow B'} [\mathbf{x}]_{B'}.$$

Main Problem

Physics: Same motion, two different observers. *Motion is relative.*
Mathematics: Same linear operator, two different bases.

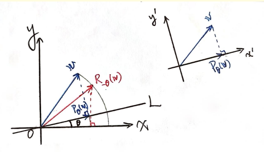
Main Problem:

V: a finite-dimensional vector space;
$B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$: a basis for V;
$B' = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\}$: another basis for V;
$T : V \to V$: a linear operator on V.

What relationship exists between the matrices $[T]_B$ and $[T]_{B'}$?

Recall the following example:

Example. Let L be the line through the origin that makes an angle θ with the positive x-axis. Find the standard matrices for the operator P_θ that maps each point into its orthogonal projection on L.



$$P_\theta = R_\theta \circ P_0 \circ R_{-\theta} = (R_{-\theta})^{-1} \circ P_0 \circ R_{-\theta}.$$

Describing a Linear Transformation

Theorem. Let $T : V \to V$ be a linear operator on a finite-dimensional vector space V, and let B and B' be bases for V. Then

$$[T]_{B'} = P^{-1}_{B\leftarrow B'}\,[T]_B\,P_{B\leftarrow B'}.$$

Proof:

$$\begin{tikzcd}
\overset{[x]_{B'}}{\mathbb{R}^n} \arrow[r, "{[T]_{B'}}"] \arrow[dd, bend right=60, "P_{B'\to B}"'] & \overset{[T(x)]_{B'}}{\mathbb{R}^n} \arrow[dd, bend left=60, "P_{B'\to B}"] \\
\overset{x}{V} \arrow[u] \arrow[d] \arrow[r, "T"] & \overset{T(x)}{V} \arrow[u] \arrow[d] \\
\underset{[x]_B}{\mathbb{R}^n} \arrow[r, "{[T]_B}"] & \underset{[T(x)]_B}{\mathbb{R}^n}
\end{tikzcd}$$

14

An Example

Example. Consider the matrix operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ whose standard matrix is

$$[T] = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}.$$

Find the matrix of T relative to another basis $B' = \{\mathbf{u}'_1, \mathbf{u}'_2\}$, in which

$$\mathbf{u}'_1 = (1, 1), \quad \mathbf{u}'_2 = (1, 2).$$

Show the relationship between this two matrices.

Solution:

15

Similar Matrices

Definition. Let $A, B \in M_n$. We say that A and B are similar, if there is an invertible matrix $P \in M_n$ such that

$$P^{-1}AP = B.$$

Remark: ◇ "Similar" is an equivalent relation on M_n, i.e.,

reflexive + symmetric + transitive properties.

◇ Similar matrices can represent the same linear operator.

Example. Verify that the matrices A and D are similar, where

$$A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

Proof:

16

Remark: The transition matrix P is not unique!

Remark: For a diagonal matrix A, one has $A\mathbf{e}_j = a_{jj}\mathbf{e}_j$.

Similarity Invariants

Theorem. Let $A, B \in M_n$ be two similar matrices. Then

- $\diamond$ $\mathrm{tr}(A) = \mathrm{tr}(B)$;
- $\diamond$ $\det(A) = \det(B)$;
- $\diamond$ $\mathrm{rank}(A) = \mathrm{rank}(B)$;
- $\diamond$ $\mathrm{nullity}(A) = \mathrm{nullity}(B)$.

Proof:

17

Remark: Trace, determinant, rank and nullity are similarity invariants.
Indeed, they characterize the underlying linear transformation.

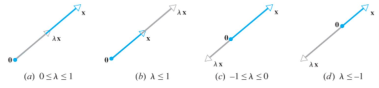
Definition of Eigenvalue & Eigenvector

Fact: A linear operator T may be very complicated!

Philosophy: Deal with some easy case first! Then simplify T!

Idea: The case that T fixes some "axis" is easy to describe.

(e.g., compression or dilation in one dimension)



Definition. Let V be an $\mathbb{F}$-vector space and let $T : V \to V$ be a linear operator. A nonzero vector $\mathbf{x} \in V$ is called an eigenvector of T if

$$T(\mathbf{x}) = \lambda \mathbf{x}$$

for some scalar $\lambda \in \mathbb{F}$. Here, the scalar λ is called an eigenvalue of T, and $\mathbf{x}$ is said to be an eigenvector corresponding to λ.

Definition of Eigenvalue & Eigenvector

Definition'. For $A \in M_n(\mathbb{F})$, if

$$A\mathbf{x} = \lambda \mathbf{x}$$

for some non-zero vector $\mathbf{x} \in \mathbb{F}^n$ and some scalar $\lambda \in \mathbb{F}$, then λ is called an eigenvalue of A, and $\mathbf{x}$ is called an eigenvector of A corresponding to λ.

Theorem. Two similar matrices have the same eigenvalues.

Proof:

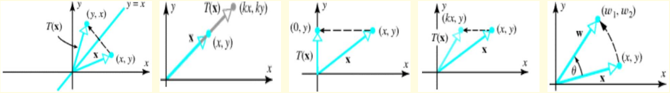
18

Remark: The corresponding eigenvectors may be different (by a transition matrix)!

Examples

Example. Find eigenvalues and eigenvectors of the following linear operators, if exist, by geometric observation.

(1) Reflection about the line $y = x$ on $\mathbb{R}^2$, i.e., $T(x, y) = (y, x)$.
(2) Contraction or dilation on $\mathbb{R}^2$, i.e., $T(x, y) = (kx, ky)$.
(3) Orthogonal projection on the y-axis in $\mathbb{R}^2$, i.e., $T(x, y) = (0, y)$.
(4) The shear of $\mathbb{R}^2$ in the x-direction, i.e., $T(x, y) = (x + ky, y)$.
(5) The rotation operator on $\mathbb{R}^2$ that moves points counterclockwise about the origin through an angle θ.



Solution:

19

Examples

Example. Let $D : C^\infty(-\infty, +\infty) \to C^\infty(-\infty, +\infty)$ be the differential operator, i.e., $D = \frac{d}{dx}$. This operator is linear. For any $\lambda \in \mathbb{R}$, the function $e^{\lambda x}$ is an eigenvector of D, i.e.,

$$D(e^{\lambda x}) = \lambda e^{\lambda x}.$$

Example. Let $T : C^\infty(-\pi, \pi) \to C^\infty(-\pi, \pi)$ be the second-order differential operator, i.e., $T = -\frac{d^2}{dx^2}$. This operator is linear. For any $k \in \mathbb{Z}$, the functions $\sin kx$ and $\cos kx$ are eigenvectors of T relative to the eigenvalue k^2, i.e.,

$$T(\sin kx) = k^2 \sin kx, \quad T(\cos kx) = k^2 \cos kx.$$

Remark: Usually, harmonic waves are eigenvectors of a Laplacian operator.

Computing Eigenvalues and Eigenvectors

- $A\mathbf{x} = \lambda \mathbf{x} \Longleftrightarrow (A - \lambda I)\mathbf{x} = \mathbf{0}$.

Theorem. Let $A \in M_n(\mathbb{F})$. Then $\lambda \in \mathbb{F}$ is an eigenvalue of A, if and only if

$$\det(\lambda I_n - A) = 0.$$

Proof:

20

Remark: "$\det(\lambda I - A) = 0$" is called the characteristic equation of A.

Definition. For $A \in M_n$, the characteristic polynomial of A is defined to be

$$f_A(\lambda) = \det(\lambda I_n - A).$$

Remark: $\deg(f_A(\lambda)) = n$. The coefficients belong to $\mathbb{F}$.

Remark: When the scalars are real, $f_A(\lambda)$ may have no root.
When the scalars are complex, $f_A(\lambda)$ has n roots (with multiplicities).

Problem*: Let $A \in M_n(\mathbb{R})$, prove that A and A^T has same eigenvalues.

Problem*: Prove that two similar matrices have the same characteristic polynomial.

Examples

Example. Let
$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

(1) Consider $A \in M_2(\mathbb{R})$, find all eigenvalues of A.
(2) Consider $A \in M_2(\mathbb{C})$, find all eigenvalues of A.

Solution:

21

Example. Find all eigenvalues of

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix} \in M_3(\mathbb{R}).$$

Solution:

22

Examples

Example. Find eigenvalues of

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}.$$

Solution:

23

Theorem. If A is a triangular matrix, then the eigenvalues of A are the entries on the main diagonal of A.

Eigenspace

Definition. Suppose that $A \in M_n(\mathbb{F})$ and λ is an eigenvalue of A. Then

$$E_\lambda(A) := \{\mathbf{x} \in \mathbb{F}^n : A\mathbf{x} = \lambda\mathbf{x}\} = \mathrm{Null}(\lambda I_n - A)$$

is called the eigenspace of A corresponding to λ.

Remark: The notation $E_\lambda(A)$ is used only in our class.
Remark: $E_\lambda(A)$ is a subspace of $\mathbb{F}^n$.
Remark: The effect of the transformation A, restricted to $E_\lambda(A)$, is a scalar multiple of λ.

Problem*: Suppose that the two matrices A and B are similar, and λ are their common eigenvalue. Prove that

$$\dim(E_\lambda(A)) = \dim(E_\lambda(B)).$$

Examples

Example. Find a basis for each eigenspace of

$$A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix} \in M_3(\mathbb{R}).$$

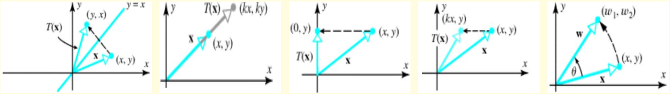
Solution:

24

Examples

Example. Find the dimension of each eigenspace of the following linear operators.

(1) Reflection about the line $y = x$ on $\mathbb{R}^2$, i.e., $T(x, y) = (y, x)$.
(2) Contraction or dilation on $\mathbb{R}^2$, i.e., $T(x, y) = (kx, ky)$.
(3) Orthogonal projection on the y-axis in $\mathbb{R}^2$, i.e., $T(x, y) = (0, y)$.
(4) The shear of $\mathbb{R}^2$ in the x-direction, i.e., $T(x, y) = (x + ky, y)$.
(5) The rotation operator on $\mathbb{R}^2$ that moves points counterclockwise about the origin through an angle θ.



Eigenvalues and Invertibility

Theorem. Let $A \in M_n(\mathbb{F})$. The following statements are equivalent.

(a) A is invertible.
(t) $\lambda = 0$ is not an eigenvalue of A.

Proof:

25

Remark: Let $p(\lambda) = \det(\lambda I_n - A)$. Then $\det(A) = (-1)^n p(0)$.
Remark: Suppose that A is invertible. If $A\mathbf{x} = \lambda \mathbf{x}$, then $A^{-1}\mathbf{x} = \lambda^{-1}\mathbf{x}$.

Let

$$A = \begin{bmatrix} 1 & 1+\mathrm{i} \\ 2 & \mathrm{i} \end{bmatrix} \in M_2(\mathbb{C}).$$

Find eigenvalues and corresponding eigenvectors of A^{-1}.
Solution:

26

Eigenvalues of Matrix Polynomials

Theorem. Let $A \in M_n(\mathbb{F})$. Let $q(x)$ be a polynomial with coefficients in $\mathbb{F}$. Then

$$\{\mu : \mu \text{ is an eigenvalue of } q(A)\} \supseteq \{q(\lambda) : \lambda \text{ is an eigenvalue of } A\}.$$

Indeed, if $A\mathbf{x} = \lambda\mathbf{x}$, then $q(A)\mathbf{x} = q(\lambda)\mathbf{x}$.

Proof:

27

Remark: When $\mathbb{F} = \mathbb{C}$, the above "$\supseteq$" can be replaced by "$=$".

Conjugate Transpose of a Matrix

Definition. Let $A \in M_{m\times n}(\mathbb{C})$, the adjoint (conjugate transpose) of A, denoted by A^*, is defined to be the matrix in $M_{n\times m}(\mathbb{C})$ such that

$$(A^*)_{ij} = \overline{(A)_{ji}}, \quad (1 \le i \le m,\ 1 \le j \le n).$$

Proposition. Let $A, B \in M_{m\times n}(\mathbb{C})$, $C \in M_{n\times k}(\mathbb{C})$ and $c \in \mathbb{C}$. Then
(i) $(A+B)^* = A^* + B^*$.
(ii) $(cA)^* = \overline{c}A^*$.
(iii) $(A^*)^* = A$.
(iv) $(AC)^* = C^*A^*$.

Theorem. Let $A \in M_n(\mathbb{C})$. If λ is an eigenvalue of A, then $\overline{\lambda}$ is an eigenvalue of A^*.

Proof:

28

Examples

Let

$$A = \begin{bmatrix} 1 & 1+\mathrm{i} \\ 2 & \mathrm{i} \end{bmatrix} \in M_2(\mathbb{C}).$$

Find the eigenvalues and corresponding eigenvectors of
(i) A^5; (ii) $A^2 + A - 3I_2$.
Find the eigenvalues of (iii) A^*.

Solution:

29

Concept of Diagonalizazion

Fact: A linear operator T may be very complicated!

Philosophy: Deal with some easy case first! Then simplify T!

Idea: The case that T fixes some "axis" is easy to describe.

Idea: Now we consider the case that T can be COMPLETELY determined by such fixed "axes"!

Example. The eigenvalues of the diagonal matrix

$$\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

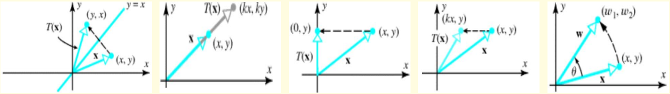
are exactly $\lambda_1, \lambda_2, \ldots, \lambda_n$. And the standard basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are the corresponding eigenvectors, respectively.

Definition. A square matrix $A \in M_n$ (or a linear operator T on V with $\dim(V) = n$) is said to be diagonalizable if it has n linearly independent eigenvectors.

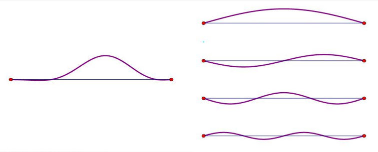
Examples

Example. Are the following linear operators diagonalizable? If so, determine the effect of the operators by the eigenvectors.

(1) Reflection about the line $y = x$ on $\mathbb{R}^2$, i.e., $T(x, y) = (y, x)$.
(2) Contraction or dilation on $\mathbb{R}^2$, i.e., $T(x, y) = (kx, ky)$.
(3) Orthogonal projection on the y-axis in $\mathbb{R}^2$, i.e., $T(x, y) = (0, y)$.
(4) The shear of $\mathbb{R}^2$ in the x-direction, i.e., $T(x, y) = (x + ky, y)$.
(5) The rotation operator on $\mathbb{R}^2$ that moves points counterclockwise about the origin through an angle θ.



Examples



Example. Let V be set of functions f on $[-\pi, \pi]$ which have derivatives of all order and satisfy $f(-\pi) = f(\pi) = 0$. The Fourier series gives

$$f(x) = \sum_{n=1}^{\infty} (a_n \sin nx + b_n \cos nx),$$

where a_n, b_n are Fourier coefficients. Indeed, $\sin nx$ and $\cos nx$ are eigenvectors of $\frac{d^2}{dx^2}$.

Remark: The theories for infinite-dimensional spaces are much more complicated. Please ignore the details!

Equivalent Statements for Diagonalizazion

Theorem. Let $A \in M_n(\mathbb{F})$. The following statements are equivalent.
(i) A has n linearly independent eigenvectors.
(ii) $P^{-1}AP = D$ for some invertible P and diagonal D.

Proof:

30

Remark: Indeed, P is the transition matrix from the basis formed by the eigenvectors to the standard basis.

Remark: Indeed, the entries on the main diagonal of D are exactly eigenvalues of A.

Diagonalise a Matrix

Examples. In $M_3(\mathbb{R})$, diagonalize A and find the corresponding transition matrix P. Here

$$A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}.$$

Solution:

31

Problem*: To summarize the main steps of diagonalization.

Remark: The choice of P is not unique!

Linearly Independent Eigenvectors

Theorem. If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of A, and $E_{\lambda_i}(A)$ $(1 \leq i \leq k)$ has basis

$$\mathbf{v}_{i,1}, \mathbf{v}_{i,2}, \ldots, \mathbf{v}_{i,m_k}.$$

Then the vectors

$$\mathbf{v}_{1,1}, \ldots, \mathbf{v}_{1,m_1}, \quad \mathbf{v}_{2,1}, \ldots, \mathbf{v}_{2,m_2}, \quad , \ldots, \quad \mathbf{v}_{k,1}, \ldots, \mathbf{v}_{k,m_k}$$

are linearly independent.

Proof:

32

Remark: Let $A \in M_n$. Suppose that all the distinct eigenvalues of A are $\lambda_1, \lambda_2, \ldots, \lambda_k$. Then A is diagonalizable if and only if

$$\dim E_{\lambda_1}(A) + \dim E_{\lambda_2}(A) + \ldots + \dim E_{\lambda_k}(A) = n.$$

Corollary. Let $A \in M_n$. If A has n distinct eigenvalues then A is diagonalizable.

Diagonalise a Matrix

Examples. Show that $B \in M_3(\mathbb{R})$ is not diagonalizable, where

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix}.$$

Solution:

33

Computing Powers of a Diagonalizable Matrix

Observation:

- If $A\mathbf{x} = \lambda\mathbf{x}$, then $A^k\mathbf{x} = \lambda^k\mathbf{x}$.
- $(P^{-1}AP)^k = P^{-1}A^kP$.
- For a polynomial $f(x)$, we have $f\left(P^{-1}AP\right) = P^{-1}f(A)P$.

Idea: Diagonalise the matrix before computing powers.

Example. Evaluate A^5, where $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$.

Solution:

34

Remark: $A^n = PD^nP$.

Recurrence Relation

Example. The Fibonacci sequence is defined by the recurrence relation

$$a_0 = 0, \quad a_1 = 1, \quad a_n = a_{n-1} + a_{n-2} \ (n \geq 2).$$

Find an expression for a_n by using 2×2 matrices.

Solution:

35