Vv417 Lecture 15

Jing Liu

UM-SJTU Joint Institute

October 29, 2019

- Recall we suspected some connection between $\mathcal{P}_2$ and $\mathbb{R}^3$

$$\underbrace{\begin{bmatrix}3\\2\\3\end{bmatrix}}_{\mathbf{u}} = 3\underbrace{\begin{bmatrix}1\\0\\1\end{bmatrix}}_{\mathbf{v}_1} + 2\underbrace{\begin{bmatrix}0\\1\\0\end{bmatrix}}_{\mathbf{v}_2} \quad \text{or} \quad \underbrace{3x^2+2x+3}_{\mathbf{u}} = 3\underbrace{(x^2+1)}_{\mathbf{v}_1} + 2\underbrace{(x)}_{\mathbf{v}_2}$$

- Specifically, the correspondence between the two vector spaces are preserved

under addition:

$$\begin{array}{ccccc} a_0 & + & a_1 x & + & a_2 x^2 \\ b_0 & + & b_1 x & + & b_2 x^2 \\ \hline (a_0+b_0) & + & (a_1+b_1)x & + & (a_2+b_2)x^2 \end{array} \iff \begin{bmatrix}a_0\\a_1\\a_2\end{bmatrix} + \begin{bmatrix}b_0\\b_1\\b_2\end{bmatrix} = \begin{bmatrix}a_0+b_0\\a_1+b_1\\a_2+b_2\end{bmatrix}$$

under scalar multiplication:

$$\beta\left(a_0+a_1x+a_2x^2\right) = \beta a_0 + (\beta a_1)x + (\beta a_2)x^2 \iff \beta\begin{bmatrix}a_0\\a_1\\a_2\end{bmatrix} = \begin{bmatrix}\beta a_0\\\beta a_1\\\beta a_2\end{bmatrix}$$

- Furthermore, it is clear that the correspondence is unique.

Q: Does it remind you of something?

- Recall that a function is a rule that associates with

  each element of a set $\mathcal{A}$ one and only one element in a set $\mathcal{B}$.

- If f associates the element b with the element a, then we write

$$b = f(a)$$

  we say that b is the image of a under f or that $f(a)$ is the value of f at a.

- The set $\mathcal{A}$ is called the domain of f and the set $\mathcal{B}$ the codomain of f.
- The subset of the codomain that consists of

  all images of elements in the domain is called the range of f.

- In many applications the domain and codomain of a function are sets of $\mathbb{R}$.
- In general, we will be concerned with functions for which the domain is a vector space $\mathcal{U}$ and the codomain is another vector space $\mathcal{V}$.

Definition

If f is a function with domain $\mathcal{U}$ and codomain $\mathcal{V}$, then we usually say that

f is a transformation from $\mathcal{U}$ to $\mathcal{V}$ or that f maps from $\mathcal{U}$ to $\mathcal{V}$,

which we denote by writing

$$f : \mathcal{U} \to \mathcal{V}$$

In the special case where $\mathcal{U} = \mathcal{V}$, it is sometimes called an operator on $\mathcal{U}$.

- It is common to use the letter T to denote a transformation. In keeping with this usage, we will usually denote a transformation from $\mathcal{U}$ to $\mathcal{V}$ by writing

$$T : \mathcal{U} \to \mathcal{V}$$

- Matrix multiplication gives a simple transformation between $\mathbb{R}^n$ and $\mathbb{R}^m$,

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

which is known as a matrix transformation, and $\mathbf{A}$ the transformation matrix

$$T_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^m$$

- Note that $\mathbb{R}^n$ is the domain of $T_{\mathbf{A}}$ but that $\mathbb{R}^m$ may not be the range of $T_{\mathbf{A}}$.

$$T_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^m$$

- The space $\mathbb{R}^m$, which is the codomain of $T_{\mathbf{A}}$, is intended only to describe the space in which the image vectors lie and may be larger than the range of $T_{\mathbf{A}}$.
- If $m = n$, $T_{\mathbf{A}}$ is known as a matrix operator, which involves a square matrix.

$$\mathbf{A}_{n \times n}$$

- Often we don't specify the domain and codomain,

$$\mathbf{y} = T_{\mathbf{A}}(\mathbf{x}) \qquad \text{or} \qquad \mathbf{x} \xrightarrow{T_{\mathbf{A}}} \mathbf{y}$$

which is read as "$T_{\mathbf{A}}$ maps $\mathbf{x}$ into $\mathbf{y}$".

Q: What does $T_{\mathbf{R}}(\mathbf{x})$ do to $\mathbf{x}$?

$$T_{\mathbf{R}} : \mathbb{R}^3 \to \mathbb{R}^3 \qquad \text{where} \qquad \mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

- Recall that if the following set $\mathcal{B}$ is basis for a vector space $\mathcal{V}$

$$\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$$

then each $\mathbf{v} \in \mathcal{V}$ can be represented as

$$\mathbf{v} = \alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + \cdots + \alpha_n \mathbf{b}_n$$

Q: Are the coefficients in this expansion unique for every $\mathbf{v} \in \mathcal{V}$.

- Suppose we have two sets of coefficients $\{\alpha_i\}$ and $\{\beta_i\}$

$$\mathbf{v} = \sum_i^n \alpha_i \mathbf{b}_i = \sum_i^n \beta_i \mathbf{b}_i$$

$$\implies \mathbf{0} = \sum_i^n (\alpha_i - \beta_i) \mathbf{b}_i$$

- This implies $(\alpha_i - \beta_i) = 0$ for each i since $\mathcal{B}$ is a linearly independent set.

Definition

Suppose $\mathbf{v} \in \mathcal{V}$, and

$$\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\} \quad \text{is a basis for the vector space } \mathcal{V},$$

The coefficients α_i in the expansion

$$\mathbf{v} = \alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + \cdots + \alpha_n \mathbf{b}_n$$

are called the coordinates of $\mathbf{v}$ with respect to $\mathcal{B}$, and $[\mathbf{v}]_{\mathcal{B}}$ denote the vector

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

which is known as the coordinate vector of $\mathbf{v}$ with respect to $\mathcal{B}$.

- Note order is important, that is, by basis, we actually meant ordered basis when we are talking about coordinates or coordinate vector.

- If no basis is explicitly mentioned, then the standard basis is assumed.
- For example, the vector $\mathbf{v} = \begin{bmatrix} 8 \\ 7 \\ 4 \end{bmatrix}$ is understood as the coordinate vector with respect to the standard basis $\mathcal{S} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, that is,

$$\mathbf{v} = \begin{bmatrix} 8 \\ 7 \\ 4 \end{bmatrix} = 8\mathbf{e}_1 + 7\mathbf{e}_2 + 4\mathbf{e}_3 = [\mathbf{v}]_{\mathcal{S}}$$

- Of course, we can talk about coordinates of a non-euclidean vector with respect to certain non-euclidean basis. For example,

$$[\mathbf{p}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2x^2 + 3x^4, \qquad \text{where} \quad \mathcal{B} = \{x^2, x^4\}$$

- In fact, some transformations are vectors in a vector space as well, therefore those transformations possess coordinates in the same way other vectors do.

- In multivariable calculus, you may have encountered transformation

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

that maps scalars to vectors or vectors to scalars,

$$\mathbf{r}(t) = \begin{bmatrix} \cos t \\ \sin t \\ t \end{bmatrix}, \qquad \text{or} \qquad f(\mathbf{x}) = f(x_1, x_2, x_3) = |\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

Q: Are there properties of a transformation

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

that can be used to determine whether T is a matrix transformation?

Q: If we discover that a transformation

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

is a matrix transformation, how can we find a transformation matrix for it?

- The following theorem and its proof will provide the answers.

Theorem

A transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is a matrix transformation if and only if the two properties below hold for all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ and for every scalar α in $\mathcal{F}$

- 1. $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$
- 2. $T(\alpha\mathbf{u}) = \alpha T(\mathbf{u})$

Proof

- Suppose T is a matrix transformation, then properties 1. and 2. are simply two basic properties of matrix multiplication, thus clearly satisfied,

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

$$\mathbf{A}(\alpha\mathbf{u}) = \alpha(\mathbf{A}\mathbf{u}) \implies T(\alpha\mathbf{u}) = \alpha T(\mathbf{u})$$

- Conversely, suppose that properties 1. and 2. hold. We need to show that there exists an $m \times n$ matrix $\mathbf{A}$ such that

$$T(\mathbf{x}) = \mathbf{A}\mathbf{x} \qquad \text{for every vector } \mathbf{x} \text{ in } \mathbb{R}^n.$$

Proof

- Use properties 1. and 2. multiple times, we have

$$T(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_r\mathbf{u}_r) = \alpha_1 T(\mathbf{u}_1) + \alpha_2 T(\mathbf{u}_2) + \cdots + \alpha_r T(\mathbf{u}_r)$$

  for all scalars α_1, α_2, …, α_r and all vectors $\mathbf{u}_1$, $\mathbf{u}_2$, …, $\mathbf{u}_r$ in $\mathbb{R}^n$.

- For every vector $\mathbf{x}$ in $\mathbb{R}^n$, we have $\mathbf{x} = [\mathbf{x}]_{\mathcal{S}}$, where $\mathcal{S} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$,

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$$

- Then the transformation matrix $\mathbf{A}$ can be shown to exist and found by

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) \\ &= x_1 T(\mathbf{e}_1) + \cdots + x_n T(\mathbf{e}_n) \\ &= \mathbf{A}\mathbf{x} \end{aligned}$$

  where $\mathbf{A} = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$. $\square$

- When the linearity conditions are satisfied by a transformation

$$T : \mathcal{U} \to \mathcal{V}$$

the transformation T is known as a linear transformation or Homomorphism.

- So every matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a linear transformation, and every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a matrix transformation.

Theorem

Suppose $T_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^m$ and $T_{\mathbf{B}} : \mathbb{R}^n \to \mathbb{R}^m$ are matrix transformations, then

$T_{\mathbf{A}}(\mathbf{x}) = T_{\mathbf{B}}(\mathbf{x})$ for very vector $\mathbf{x}$ in $\mathbb{R}^n$ if and only if $\mathbf{A} = \mathbf{B}$

- The last theorem is significant because it tells us that there is a

one-to-one correspondence

between $m \times n$ matrices and matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.

- So every $m \times n$ matrix produces exactly 1 matrix transformation, every matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ has exactly 1 transformation matrix.

Definition

The matrix with the image vectors of the standard vectors as its columns

$$\begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$$

is called the standard matrix for the transformation.

Exercise

Find the standard matrix $\mathbf{R}_\theta$ for the rotation operator

$$T_\theta \colon \mathbb{R}^2 \to \mathbb{R}^2$$

that moves points counterclockwise about the origin through a positive angle θ.

Solution

- The image vector of $\mathbf{e}_1$ and $\mathbf{e}_2$ under T_θ are the columns of $\mathbf{R}_\theta$.

- Suppose that

$$T_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^k, \qquad \text{and} \qquad T_{\mathbf{B}} : \mathbb{R}^k \to \mathbb{R}^m$$

  are linear transformations.

- Consider $\mathbf{x} \in \mathbb{R}^n$, then

$$T_{\mathbf{A}}$$

  maps this vector into a vector $T_{\mathbf{A}}(\mathbf{x})$ in $\mathbb{R}^k$, and

$$T_{\mathbf{B}}$$

  in turn, maps that vector into the vector $T_{\mathbf{B}}(T_{\mathbf{A}}(\mathbf{x}))$ in $\mathbb{R}^m$.

- Together this creates a transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ that we call

  the composition of $T_{\mathbf{B}}$ with $T_{\mathbf{A}}$ and denote by the symbol $T_{\mathbf{B}} \circ T_{\mathbf{A}}$

- The transformation $T_{\mathbf{A}}$ in the formula is performed first; that is,

$$(T_{\mathbf{B}} \circ T_{\mathbf{A}})(\mathbf{x}) = T_{\mathbf{B}}(T_{\mathbf{A}}(\mathbf{x})) = \mathbf{BAx}$$

- Recall matrix multiplications are Not commutative

$$\mathbf{AB} \neq \mathbf{BA}$$

- So compositions of linear transformations are Not commutative in general

$$[T_2 \circ T_1] \neq [T_1 \circ T_2]$$

- For example, if $T_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^k$ and if $T_{\mathbf{B}} : \mathbb{R}^k \to \mathbb{R}^m$, then

$$(T_{\mathbf{B}} \circ T_{\mathbf{A}})(\mathbf{x}) = \mathbf{BAx} \neq \mathbf{BAx} = (T_{\mathbf{A}} \circ T_{\mathbf{B}})(\mathbf{x})$$

Q: Is the composition of rotations commutative?

$$\mathbf{A} = \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix}$$

- Since

$$\mathbf{AB} = \mathbf{BA}$$

in this case, but in general linear transformations are not commutative.

- Our next objective is to establish a link between the invertibility of a matrix $\mathbf{A}$ and properties of the corresponding matrix transformation $T_{\mathbf{A}}$.

Definition

A matrix transformation $T_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^m$ is said to be one-to-one if $T_{\mathbf{A}}$ maps distinct vectors in $\mathbb{R}^n$ into distinct vectors in $\mathbb{R}^m$.

Q: Are rotation operators on $\mathbb{R}^2$ one-to-one?

Q: How about the following transformation?

$$T_{\mathbf{A}}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ 2x_1 + 5x_2 \\ 3x_1 + 4x_2 \end{bmatrix} = x_1 \overset{\mathbf{u}}{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}} + x_2 \overset{\mathbf{v}}{\begin{bmatrix} -1 \\ 5 \\ 4 \end{bmatrix}}$$

- It can easily shown that $\mathbf{u}$ and $\mathbf{v}$ are linearly independent, thus for distinct $\mathbf{x}$

$$T_{\mathbf{A}}(\mathbf{x}) - T_{\mathbf{A}}(\mathbf{x}^*) = x_1\mathbf{u} + x_2\mathbf{v} - (x_1^*\mathbf{u} + x_2^*\mathbf{v}) = (x_1 - x_1^*)\mathbf{u} + (x_2 - x_2^*)\mathbf{v} \neq 0$$

Definition

Let $T_{\mathbf{A}}: \mathbb{R}^n \to \mathbb{R}^m$ be a matrix transformation, then the set of all vectors in $\mathbb{R}^n$ that $T_{\mathbf{A}}$ maps into $\mathbf{0}$ is called the kernel of $T_{\mathbf{A}}$ and is denoted by

$$\text{kernel}(T_{\mathbf{A}}) = \text{null}(\mathbf{A})$$

The set of all image vectors in $\mathbb{R}^m$ under the transformation is called the range of $T_{\mathbf{A}}$ and is denoted by

$$\text{range}(T_{\mathbf{A}}) = \text{col}(\mathbf{A})$$

- The kernel and the range are in terms of the transformation matrix $\mathbf{A}$.

- Matrix view: $\text{null}(\mathbf{A})$
- System view: the solution space of

$$\mathbf{Ax} = \mathbf{0}$$

- Transformation view: $\text{kernel}(T_{\mathbf{A}})$

- Matrix view: $\text{col}(\mathbf{A})$
- System view: all $\mathbf{b}$ for which

$$\mathbf{Ax} = \mathbf{b} \quad \text{is consistent}$$

- Transformation view: $\text{range}(T_{\mathbf{A}})$

Theorem

A matrix transformation $T_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one if and only if

$$\text{kernel}(T_{\mathbf{A}}) = \{\mathbf{0}\}$$

Proof

- Assume one-to-one, then the trivial solution is the only solution

$$\mathbf{A}\mathbf{x} = \mathbf{0}$$

  so the null space and thus the kernel have only the zero vector.

- Assume there is only the zero vector in the kernel of $T_{\mathbf{A}}$, then

$$\mathbf{A}\mathbf{x} = \mathbf{0}$$

  has only the trivial solution. So the columns of $\mathbf{A}$ are linearly independent.

- Let $\mathbf{x}$ and $\mathbf{y}$ be two distinct solutions, then their images are distinct, so 1-1.

$$\mathbf{A}\mathbf{x} - \mathbf{A}\mathbf{y} = (x_1 - y_1)\mathbf{a}_1 + (x_1 - y_1)\mathbf{a}_2 + \cdots (x_1 - y_1)\mathbf{a}_n \neq 0 \quad \square$$

Definition

The matrix operator that corresponds to $\mathbf{A}^{-1}$

$$T_{\mathbf{A}^{-1}} : \mathbb{R}^n \to \mathbb{R}^n$$

is called the inverse operator, or more simply, the inverse of $T_{\mathbf{A}}$.

Exercise

Find the inverse of the rotation operator $\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ on $\mathbb{R}^2$.

Solution

- We can find the inverse of the transformation matrix $\mathbf{A}^{-1}$ by elimination.
- However, it is clear geometrically that to undo the effect of rotation by θ, we only need to rotate each vector through $-\theta$,

$$\mathbf{R}^{-1} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Equivalence Theorem

If $\mathbf{A}$ is an $n \times n$ matrix, then the following statements are equivalent,

1. $\mathbf{A}$ is invertible.
2. $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.
3. The reduced echelon form of $\mathbf{A}$ is $\mathbf{I}_n$.
4. $\mathbf{A}$ is expressible as a product of elementary matrices.
5. $\mathbf{Ax} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.
6. $\mathbf{Ax} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.
7. $\det(\mathbf{A}) \neq 0$.
8. The column vectors of $\mathbf{A}$ are linearly independent.
9. The row vectors of $\mathbf{A}$ are linearly independent.
10. $\mathbf{A}$ has rank n
11. $\mathbf{A}$ has nullity 0.
12. The kernel of T_A is $\{\mathbf{0}\}$.
13. The range of T_A is $\mathbb{R}^n$.
14. T_A is one-to-one.