线性代数 (Linear Algebra)

第五章 Eigenvalues and Eigenvectors

§ 5.4 Eigenvectors and Linear Transformations
特征向量和线性变换

衡 益

2021 年 12 月 16 日，中山大学南校区

学习目标

在本节，我们将矩阵分解

$$A = PDP^{-1}$$

理解为线性变换。我们还将看到变换$x \mapsto Ax$ 实质上是简单的映射$u \mapsto Du$. 即使D不是对角矩阵，对矩阵A 和D 仍有相似的解释。

之前的内容中曾讲到，任意一个从$\mathbb{R}^n$ 到$\mathbb{R}^m$ 的线性变换T 可通过左乘矩阵A来实现，矩阵A 称为T 的标准矩阵。现在，我们对两个有限维向量空间之间的线性变换也作同样的描述。

要点一：线性变换的矩阵表示

要点二：从线性变换的角度来看$A = PDP^{-1}$，意义?

2

线性变换矩阵

Review

定理（唯一表示定理）

令B=$\{\mathbf{b}_1, \cdots, \mathbf{b}_n\}$是向量空间 V 的一个基，

则对 V 中每个向量x, 存在唯一的一组数$c_1, \cdots, c_n$,

使得：$\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$

$$\mathbb{R}^n$$

Review

定义 假设集合B={ $\mathbf{b}_1, \cdots, \mathbf{b}_n$ }是 V的一个基，x在 V中，x相对于基B的坐标(或x的B-坐标)是使得 $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$的权 $c_1, \cdots, c_n$.

若 $c_1, \cdots, c_n$是x的B-坐标，则$\mathbb{R}^n$中的向量

$$[\mathbf{x}]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

是x(相对于B)的坐标向量，

映射 $\mathbf{x} \to [\mathbf{x}]_B$ 称为（由B确定的）坐标映射.

Review

$\mathbb{R}^n$中的坐标

Review

- Example : Let $b_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, b_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, x = \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \mathcal{B} = \{b_1, b_2\}$

  Find the coordinate vector $\boxed{[\mathbf{x}]_\beta \text{ of } \mathbf{x} \text{ relative to } \boldsymbol{\beta}}$

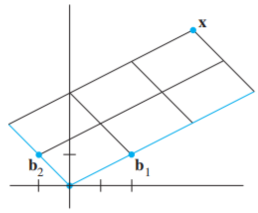
- Solution : The $\boldsymbol{\beta}$-coordinate c_1, c_2 of $\mathbf{x}$ satisfy

$$\underset{b_1}{c_1\begin{bmatrix} 2 \\ 1 \end{bmatrix}} + \underset{b_2}{c_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}} = \underset{x}{\begin{bmatrix} 4 \\ 5 \end{bmatrix}} \quad \text{or} \quad \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$$\boxed{b_1 \quad b_2 \qquad\qquad x}$$

So $c_1 = 3, c_2 = 2$

$$x = 3b_1 + 2b_2 \qquad \boxed{[x]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}}$$



7

Review

For a basis $\beta = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, let

$$P_\beta = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix} \text{ and } [\mathbf{x}]_\beta = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

Then

$$\mathbf{x} = P_\beta[\mathbf{x}]_\beta.$$

We call P_β the **change-of-coordinates matrix** from β to the standard basis in $\mathbf{R}^n$. Then

回到标准坐标

$$[\mathbf{x}]_\beta = P_\beta^{-1}\mathbf{x}$$

and therefore P_β^{-1} is a **change-of-coordinates matrix** from the standard basis in $\mathbf{R}^n$ to the basis β.

8

线性变换矩阵

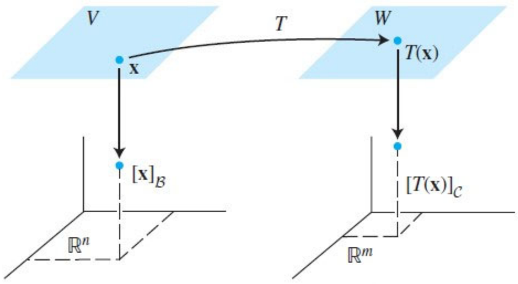


FIGURE 1 A linear transformation from V to W.

V：n 维向量空间

W：m 维向量空间

T：从V到W的线性变换

9

线性变换矩阵

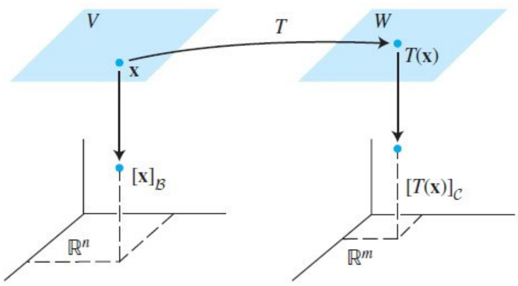


FIGURE 1 A linear transformation from V to W.

B：V空间的一组基

C：W空间的一组基

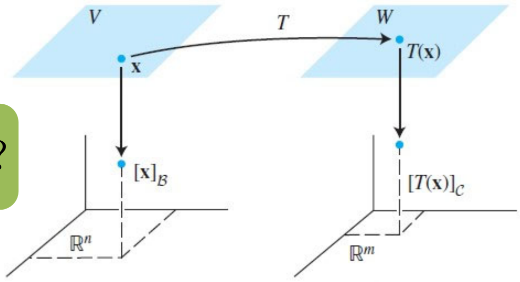
$[\mathbf{x}]_B$：$\mathbf{x} \in V$在$\mathbb{R}^n$中的坐标向量

$[T(\mathbf{x})]_C$：x的像$T(\mathbf{x})$在$\mathbb{R}^m$中的坐标向量

10

线性变换矩阵

$[\mathrm{x}]_{\mathrm{B}}$ 关系 $[T(\mathrm{x})]_{\mathrm{C}}$?



设B $=\{\mathrm{b}_1,\cdots,\mathrm{b}_n\}$，是空间V 的一组基。若

$$\mathrm{x}=r_1\mathrm{b}_1+\cdots+r_n\mathrm{b}_n,\quad r_1,\ldots,r_n\in\mathbb{R}$$

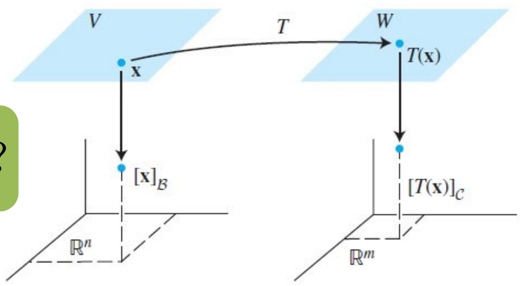
那么，

$$[\mathrm{x}]_{\mathrm{B}}=\begin{bmatrix} r_1\\ \vdots\\ r_n\end{bmatrix}$$

11

线性变换矩阵

$[\mathrm{x}]_{\mathrm{B}}$ 关系 $[T(\mathrm{x})]_{\mathrm{C}}$?



由于T是线性变换，则

$$T(\mathrm{x})=T(r_1\mathrm{b}_1+\cdots+r_n\mathrm{b}_n)=r_1T(\mathrm{b}_n)+\cdots+r_nT(\mathrm{b}_n)$$

因为坐标变换是线性变换，则

$$[T(\mathrm{x})]_{\mathrm{C}}=r_1[T(\mathrm{b}_1)]_{\mathrm{C}}+\cdots+r_n[T(\mathrm{b}_n)]_{\mathrm{C}}$$

由于C的坐标向量都属于$\mathbb{R}^m$，那么上式可以写成

$$[T(\mathrm{x})]_{\mathrm{C}}=\mathrm{M}[\mathrm{x}]_{\mathrm{B}}\qquad(\#)$$

12

线性变换矩阵

$[x]_B$ 关系 $[T(x)]_C$?

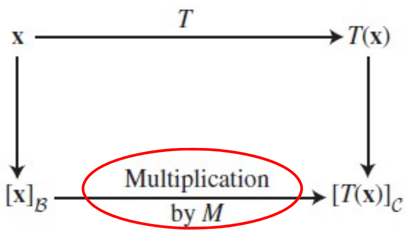


FIGURE 2

其中

$$M = \left[[T(b_1)]_c \quad [T(b_2)]_c \quad \cdots \quad [T(b_n)]_c \right]$$

矩阵M是T 的矩阵表示，称为T 相对于基B 和C 的矩阵，见图2. 等式$[T(x)]_C = M[x]_B$ 表明，就坐标向量而言，T 对x 的作用相当于用矩阵M 左乘x.

13

线性变换矩阵

例1 设$B = \{b_1, b_2\}$是V的一组基，$C = \{c_1, c_2, c_3\}$是W 的一组基。设T: V $\to$ W是线性变换，满足:

$$T(b_1) = 3c_1 - 2c_2 + 5c_3 \qquad T(b_2) = 4c_1 + 7c_2 - c_3$$

写出T 相对于B和C的矩阵M .

解 b_1和b_2像的C坐标向量为:

$$[T(b_1)]_C = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix}, \qquad [T(b_2)]_C = \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}$$

因此

$$M = \begin{bmatrix} 3 & 4 \\ -2 & 7 \\ 5 & -1 \end{bmatrix}$$

14

$V \rightarrow V$的线性变换

V→V的线性变换

$$
\begin{tikzcd}
\mathbf{x} \arrow[r, "T"] \arrow[d] & T(\mathbf{x}) \arrow[d] \\
{[\mathbf{x}]_{\mathcal{B}}} \arrow[r, "\text{Multiplication}", "\text{by } {[T]_{\mathcal{B}}}"'] & {[T(\mathbf{x})]_{\mathcal{B}}}
\end{tikzcd}
$$

若取W ＝ V，C ＝ B时，那么矩阵M 称作：线性变换T 相对于B的矩阵，记作$\left[T\right]_{B}$，或简称为T 的B－矩阵。V $\mapsto$ V 的线性变换T 的B－矩阵对于所有V 中的x ，有

$$\left[T(\mathrm{x})\right]_{B} = \left[T\right]_{B}\left[\mathrm{x}\right]_{B}.$$

举例

例2 映射T：$P_2 \to P_2$定义为：

$$T(a_0 + a_1 t + a_2 t^2) = a_1 + 2a_2 t$$

映射T是线性变换。

(1)写出T在B下的矩阵，$B = \{1, t, t^2\}$.

(2)验证，对于每个$p \in P_2$，$[T(p)]_B = [T]_B [p]_B$.

解 (1) 计算基向量的像：

$T(1) = 0$　　零多项式

$T(t) = 1$　　始终为1的多项式

$T(t^2) = 2t$

17

举例

例2 映射T：$P_2 \to P_2$定义为：

$$T(a_0 + a_1 t + a_2 t^2) = a_1 + 2a_2 t$$

映射T是线性变换。

(1)写出T在B下的矩阵，$B = \{1, t, t^2\}$.

(2)验证，对于每个$p \in P_2$，$[T(p)]_B = [T]_B [p]_B$.

解

写出$T(1)$, $T(t)$和$T(t^2)$ 的B坐标向量，并将他们放在一起作为T 的B 矩阵：

$$[T(1)]_B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad [T(t)]_B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [T(t^2)]_B = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

$$[T]_B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

18

举例

例2 映射T：$P_2 \to P_2$定义为：

$$T(a_0 + a_1 t + a_2 t^2) = a_1 + 2a_2 t$$

映射T是线性变换。

(1)写出T在B下的矩阵，$B = \{1, t, t^2\}$.

(2)验证，对于每个$p \in P_2$，$[T(p)]_B = [T]_B [p]_B$.

解

(2) 对于一般的$p(t) = a_0 + a_1 t + a_2 t^2$

$$[T(p)]_B = [a_1 + 2a_2 t]_B = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = [T]_B [p]_B$$

19

注解

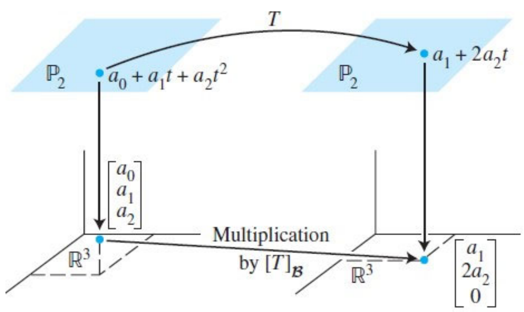
例2 映射T：$P_2 \to P_2$定义为：

$$T(a_0 + a_1 t + a_2 t^2) = a_1 + 2a_2 t$$

映射T是线性变换。

(1)写出T在B下的矩阵，$B = \{1, t, t^2\}$.

(2)验证，对于每个$p \in P_2$，$[T(p)]_B = [T]_B [p]_B$.



20

R^n 上的线性变换

21

R^n上的线性变换

在涉及$\mathbb{R}^n$ 的应用问题中，线性变换首先表现为一个矩阵变换

$$x \mapsto Ax$$

假设A是可对角化的，那么存在由A 的特征向量组成的$\mathbb{R}^n$的基。此时，下面的定理表明T 的B－矩阵是对角矩阵。

这样，把A对角化相当于找到变换$x \mapsto Ax$ 的对角矩阵。

22

$\mathbf{R}^n$上的线性变换

定理 对角矩阵的表示

设$A = PDP^{-1}$，其中D是n 阶对角矩阵。若B 是$\mathbb{R}^n$ 的一组基，且B是由P 中的列向量组成的，那么D 是线性变换$x \to Ax$的B-矩阵。

证明

设矩阵P 的列向量为$b_1, \cdots, b_n$, 那么$B = \{b_1, \cdots, b_n\}$，且$P = [b_1, \cdots, b_n]$. 那么P是坐标变换矩阵P_B ，其中

$$P[x]_B = x \quad 和 \quad [x]_B = P^{-1}x$$

$\mathbf{R}^n$上的线性变换

定理 对角矩阵的表示

设$A = PDP^{-1}$，其中D是n 阶对角矩阵。若B 是$\mathbb{R}^n$ 的一组基，且B是由P 中的列向量组成的，那么D 是线性变换$x \to Ax$的B-矩阵。

证明 若$T(x) = Ax, x \in \mathbb{R}^n$, 那么

$$\begin{aligned}
[T]_B &= \left[[T(b_1)]_B \cdots [T(b_n)]_B\right] && [T]_B \text{ 的定义} \\
&= \left[[Ab_1]_B \cdots [Ab_1]_B\right] && \text{由于} T(x) = Ax \qquad (*) \\
&= \left[P^{-1}Ab_1 \cdots P^{-1}Ab_n\right] && \text{坐标变换} \\
&= P^{-1}A\left[b_1 \cdots b_n\right] && \text{矩阵乘法} \\
&= P^{-1}AP
\end{aligned}$$

由于$A = PDP^{-1}$，那么我们可得$[T]_B = P^{-1}AP = D$

R^n上的线性变换

例3 设T: $\mathbb{R}^2 \to \mathbb{R}^2$, $T(x) = Ax$，其中$A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. 找出$\mathbb{R}^2$ 的一组基B，使得T 的B-矩阵是对角阵。

解 $A = PDP^{-1}$，其中

$$P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}, \qquad D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

P的列向量b_1，b_2是A 的特征向量。由定理 可知D 是T在$B = \{b_1, b_2\}$下的矩阵。

映射$x \mapsto Ax$ 和$u \mapsto Du$ 描述了不同基下相同的线性变换。

矩阵表示的相似性

矩阵表示的相似性

$$\begin{tikzcd}[column sep=huge, row sep=large]
\mathbf{x} \arrow[r, "\text{Multiplication}", "\text{by } A"'] \arrow[d, "\text{Multiplication}"', "\text{by } P^{-1}"] & A\mathbf{x} \\
{[\mathbf{x}]_{\mathcal{B}}} \arrow[r, "\text{Multiplication}", "\text{by } C"'] & {[A\mathbf{x}]_{\mathcal{B}}} \arrow[u, "\text{Multiplication}"', "\text{by } P"]
\end{tikzcd}$$

若A相似于C，即

$$A = PCP^{-1}$$

那么当B 是由P 的列向量构成时，C 是线性变换 $x \mapsto Ax$的B-矩阵。

矩阵表示的相似性

$$\begin{tikzcd}[column sep=huge, row sep=large]
\mathbf{x} \arrow[r, "\text{Multiplication}", "\text{by } A"'] \arrow[d, "\text{Multiplication}"', "\text{by } P^{-1}"] & A\mathbf{x} \\
{[\mathbf{x}]_{\mathcal{B}}} \arrow[r, "\text{Multiplication}", "\text{by } C"'] & {[A\mathbf{x}]_{\mathcal{B}}} \arrow[u, "\text{Multiplication}"', "\text{by } P"]
\end{tikzcd}$$

相反，若$T: \mathbb{R}^n \mapsto \mathbb{R}^n$ 由

$$T(x) = Ax$$

定义，并且若B是$\mathbb{R}^n$ 的任意一组基，那么T 的B－矩阵与A 相似。

其实，在（＊）的计算中已经证明，若P 是以B 的向量作为列构成的矩阵，那么$[T]_B = P^{-1}AP$. 因此，所有相似于A 的矩阵的集合与变换$x \mapsto Ax$ 的所有矩阵表示的集合是同一集合。

举例

例4 设$A=\begin{bmatrix}4&-9\\4&-8\end{bmatrix}, b_1=\begin{bmatrix}3\\2\end{bmatrix}, b_2=\begin{bmatrix}2\\1\end{bmatrix}$. A的特征多项式是$(\lambda+2)^2$，但特征值-2所对应的特征空间的维数为1 ，因此A 不可对角化。令$B=\{b_1,b_2\}$为一组基，线性变换$x\mapsto Ax$ 的B－矩阵是三角矩阵，称作若尔当矩阵，写出此B－矩阵。

解 若$P=\begin{bmatrix}b_1 & b_2\end{bmatrix}$，那么线性变换的B－矩阵是$P^{-1}AP$，计算可得

$$AP=\begin{bmatrix}4&-9\\4&-8\end{bmatrix}\begin{bmatrix}3&2\\2&1\end{bmatrix}=\begin{bmatrix}-6&-1\\-4&0\end{bmatrix}$$

$$P^{-1}AP=\begin{bmatrix}-1&2\\2&-3\end{bmatrix}\begin{bmatrix}-6&-1\\-4&0\end{bmatrix}=\begin{bmatrix}-2&1\\0&-2\end{bmatrix}$$

注意到A 的特征值为对角线上的元素。

29

Remark

数值计算注解

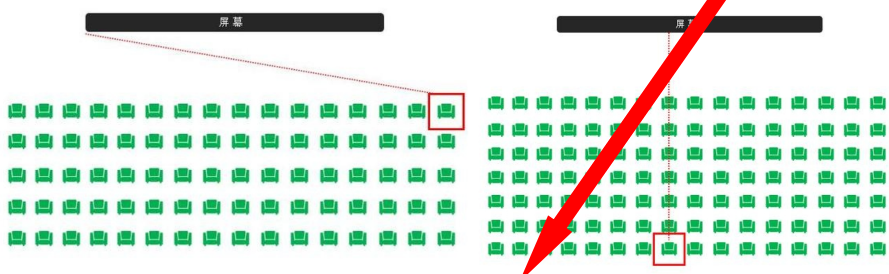
计算B－矩阵$P^{-1}AP$ 的一个有效的方法是，先计算AP ，然后利用行变换将增广矩阵[P AP] 化为$[I\ P^{-1}AP]$. 这样就不需要单独计算P^{-1} 了！

30

Remark

相似性vs线性变换

回顾刚才看电影的例子，同一部电影，在不同的位置会有不同的效果，因此“电影”是不变的，那么“电影”实质上是什么呢？



线性变换！

31

Q & A