教　案

线形变换及其矩阵表示

教学内容

线形变换是几何空间和函数空间中最简单的变换，它有着深刻的几何学和物理学背景，是一个经常使用的数学工具，在数学理论研究和实际应用中起着重要作用。在这节中主要讲解以下几方面的内容：

（1）　线性变换的概念、乘积变换和可逆变换的概念；

（2）　线性变换的矩阵表示；

（3）　在不同基下的表示矩阵之间的关系；

（4）　在线性变换下坐标的变化情况。

教学思路和要求

（1）　线性空间与线性变换这部分内容，由于其抽象性较强，所以首先要把其背景讲清楚，再抽象其理论定义。因此在教学安排上先从简单的几何变换入手，引导学生理解线性变换的概念的深刻含义，以及应有的形式。

（2）　线性变换的概念、线性变换的矩阵表示是本节内容的重点；

（3）　线性变换在不同基下的表示矩阵之间的关系也是要求必须掌握的内容；

（4）　为了使学生理解线性变换的矩阵表示的方法，可以先从向量空间上入手，便于理解；

（5）　要通过例子来引导学生学会计算表示矩阵以及不同基下的表示矩阵之间的关系等。

教学安排

一．几个简单的几何变换

复杂的几何变换可以归结为简单的几何变换的累积，而任何几何图形的变换，说到底是点的变换。

我们先从 $\mathbf{R}^2$ 谈起。容易发现，若给定了一个 2×2 矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

则对平面上任意点（即向量）$\boldsymbol{x}=\begin{pmatrix} x \\ y \end{pmatrix}$，通过矩阵与向量的乘法运算

$$\boldsymbol{x}'=A\boldsymbol{x}=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} x' \\ y' \end{pmatrix},$$

可以唯一确定了平面上的一点 $\boldsymbol{x}'$。$\boldsymbol{x}'$ 可以看成是由 $\boldsymbol{x}$ 经过某种变换得到的点，而这个变换的规律显然由矩阵 A 所确定。

例 5.2.1　问以下矩阵对 $\mathbf{R}^2$ 上的任意点 $\boldsymbol{x}$，由 $\boldsymbol{x}'=A_i\boldsymbol{x}$（$i=1,2,3,4,5$）确定了什么样的变换？

（1）$A_1=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$；　　（2）$A_2=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$；　　（3）$A_3=\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$；

（4）$A_4=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$；　　（5）$A_5=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$。

解　（1）由于对任意点 $\boldsymbol{x}=\begin{pmatrix}x\\ y\end{pmatrix}$，有

$$\boldsymbol{x}'=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}x\\ -y\end{pmatrix},$$

所以 A_1 确定的变换将任意一个点 $\boldsymbol{x}$ 变成它关于 x 轴对称的点 $\boldsymbol{x}'$（见图 5.2.1）。

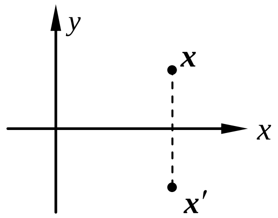

图 5.2.1

（2）由于对任意点 $\boldsymbol{x}=\begin{pmatrix}x\\ y\end{pmatrix}$，有

$$\boldsymbol{x}'=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}y\\ x\end{pmatrix},$$

所以 A_2 确定的变换将任意一个点 $\boldsymbol{x}$ 变成它关于直线 $y=x$ 对称的点 $\boldsymbol{x}'$（见图 5.2.2）。

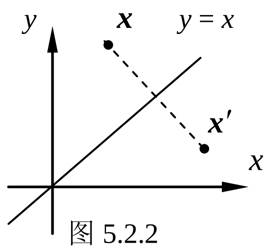

图 5.2.2

（3）由于对任意点 $\boldsymbol{x}=\begin{pmatrix}x\\ y\end{pmatrix}$，有

$$\boldsymbol{x}'=\begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}\lambda x\\ \lambda y\end{pmatrix},$$

所以 A_3 确定的变换将任意一个点 $\boldsymbol{x}$ 变成在它与原点连线上，与原点距离伸缩为 $|\lambda|$ 倍的点 $\boldsymbol{x}'$，当 $\lambda>0$ 时，$\boldsymbol{x}'$ 与 $\boldsymbol{x}$ 在原点同侧；当 $\lambda<0$ 时，$\boldsymbol{x}'$ 点在原点另一侧；当 $\lambda=0$，$\boldsymbol{x}'$ 为原点（见图 5.2.3）。

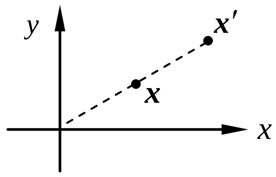

图 5.2.3

（4）对任意点 $\boldsymbol{x}=\begin{pmatrix}x\\ y\end{pmatrix}$，将其记为 $\begin{pmatrix}r\cos\varphi\\ r\sin\varphi\end{pmatrix}$，则有

$$\begin{aligned}\boldsymbol{x}'=A_4\begin{pmatrix}x\\ y\end{pmatrix}&=\boldsymbol{x}'\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}r\cos\varphi\\ r\sin\varphi\end{pmatrix}\\ &=\begin{pmatrix}r(\cos\theta\cos\varphi-\sin\theta\sin\varphi)\\ r(\sin\theta\cos\varphi+\cos\theta\sin\varphi)\end{pmatrix}\\ &=\begin{pmatrix}r\cos(\varphi+\theta)\\ r\sin(\varphi+\theta)\end{pmatrix},\end{aligned}$$

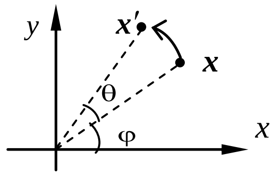

图 5.2.4

所以 A_4 确定的变换将任意一个点 $\boldsymbol{x}$ 绕原点旋转了角度 θ 的点 $\boldsymbol{x}'$（见图 5.2.4）。

（5）由于对任意点 $\boldsymbol{x}=\begin{pmatrix}x\\ y\end{pmatrix}$，有

$$\boldsymbol{x}'=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}x\\ 0\end{pmatrix},$$

所以 A_5 确定的变换将任意一个点 $\boldsymbol{x}$ 变成它在 x 轴上的投影点 $\boldsymbol{x}'$（见图 5.2.5）。

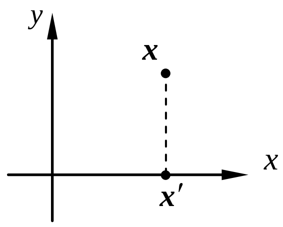

图 5.2.5

在上面的讨论中，变换由矩阵 A 确定，因此称 A

为**变换矩阵**。其中，A_1与A_2确定的变换称为**反射变换**或**镜像变换**，A_3确定的变换称为**相似变换**（λ称为**相似比**），而A_4确定的变换称为**旋转变换**，A_5确定的变换称为**射影变换**，它们都属于最简单的几何变换。

从这几个具体例子容易归纳出：

（1）设$\boldsymbol{x}_1$和$\boldsymbol{x}_2$都是平面上的点，若对它们的线性组合$\alpha_1\boldsymbol{x}_1+\alpha_2\boldsymbol{x}_2$作上述变换，可以先对$\boldsymbol{x}_1$和$\boldsymbol{x}_2$作上述变换后再线性组合，即

$$A_i(\alpha_1\boldsymbol{x}_1+\alpha_2\boldsymbol{x}_2)=\alpha_1(A_i\boldsymbol{x}_1)+\alpha_2(A_i\boldsymbol{x}_2)。$$

也就是说，由矩阵确定的变换都满足线性运算规则。

（2）如需要先将$\boldsymbol{x}$关于直线$y=x$作对称，再旋转角度θ，则有

$$\boldsymbol{x}'=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\left[\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\boldsymbol{x}\right]=\left[\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\right]\boldsymbol{x},$$

也就是说，由矩阵确定的变换可以复合，复合的变换矩阵恰是各个变换矩阵的乘积。

（3）有些变换可以通过相反的过程再变换回去，即变换是可逆的，有些则不可逆。如上面由A_1—A_4确定的变换都是可逆的，而A_5确定的变换不可逆。而通过观察发现，恰恰A_1—A_4都是可逆矩阵，而A_5是不可逆矩阵。因而可以设想，若矩阵A不可逆，那么A确定的变换不可逆；若A可逆，那么A确定的变换可逆，且确定逆变换的矩阵正是A^{-1}。

显然，借助矩阵会给讨论问题带来很大方便。于是自然要问，既然有一个矩阵就决定了一个变换，那么什么样的变换才可以通过矩阵来表示？进一步，这样的变换有哪些更一般的性质？下面来回答这些问题。

二．线性变换及其矩阵表示

定义 5.2.1　设$\mathbf{U}$，$\mathbf{V}$是$\mathbf{K}$上的线性空间，$\mathbf{K}$为$\mathbf{R}$或$\mathbf{C}$，$\boldsymbol{A}$是$\mathbf{U}$到$\mathbf{V}$的映射，即对于任意$\boldsymbol{x}\in\mathbf{U}$，存在唯一的像$\boldsymbol{z}\in\mathbf{V}$，使得$\boldsymbol{A}(\boldsymbol{x})=\boldsymbol{z}$。

若$\boldsymbol{A}$满足线性性质，即对于任意$\boldsymbol{x}$，$\boldsymbol{y}\in\mathbf{U}$及$\lambda$，$\mu\in\mathbf{K}$，成立

$$\boldsymbol{A}(\lambda\boldsymbol{x}+\mu\boldsymbol{y})=\lambda\boldsymbol{A}(\boldsymbol{x})+\mu\boldsymbol{A}(\boldsymbol{y}),$$

则称$\boldsymbol{A}$为线性空间$\mathbf{U}$到$\mathbf{V}$上的一个**线性变换**。

特别地，从线性空间$\mathbf{U}$到其自身的线性变换称为$\mathbf{U}$上的线性变换。

显然，例 5.2.1 中的五个变换都是$\mathbf{R}^2$上的线性变换。

几个最简单的线性变换是：

（1）线性空间$\mathbf{U}$上的**恒等变换**（**单位变换**）$\boldsymbol{I}$：对于任意$\boldsymbol{x}\in\mathbf{U}$，$\boldsymbol{I}(\boldsymbol{x})=\boldsymbol{x}$。

（2）线性空间$\mathbf{U}$到$\mathbf{V}$上的**零变换**$\boldsymbol{0}$：对于任意$\boldsymbol{x}\in\mathbf{U}$，$\boldsymbol{0}(\boldsymbol{x})=\mathbf{0}$。

例 5.2.2　证明求导运算$\boldsymbol{D}=\dfrac{d}{dx}$是$\mathbf{P}_n$的上的线性变换。

证　对与$\mathbf{P}_n$中的任意元素$p=p(x)$，$p(x)$是不超过n次的多项式，于是$\boldsymbol{D}(p)$ $=\dfrac{d}{dx}[p(x)]$是不超过$n-1$次的多项式，即$\boldsymbol{D}(p)\in\mathbf{P}_n$。

对于任意$p(x)$，$q(x)\in\mathbf{P}_n$及λ，$\mu\in\mathbf{R}$，由求导运算法则，

$$\boldsymbol{D}(\lambda p+\mu q)=\frac{d}{dx}(\lambda\ p(x)+\mu q(x))$$

$$= \lambda \frac{d}{dx}[\, p(x)\,] + \mu \frac{d}{dx}[\, q(x)\,] = \lambda \boldsymbol{D}(p) + \mu \boldsymbol{D}(q),$$

由定义，$\boldsymbol{D}$ 是 $\mathbf{P}_n$ 上的线性变换。

例 5.2.3 求定积分运算 $\boldsymbol{L}(f) = \int_a^b f(x)dx$ 是 $C[a,b]$ 到 $\mathbf{R}$ 上的线性变换。实际上，$\boldsymbol{L}$ 的线性性质就是定积分的线性性质。

例 5.2.4 设映射 $\boldsymbol{A}$： $\mathbf{R}^3 \to \mathbf{R}^3$ 定义为

$$\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{A}(x_1, x_2, x_3) = \left(\sqrt{x_1^2 + x_2^2}, \sqrt{x_2^2 + x_3^2}, \sqrt{x_3^2 + x_1^2}\right)^T 。$$

则对于 $\lambda \in \mathbf{R}$，有

$$\boldsymbol{A}(\lambda \boldsymbol{x}) = \left(\sqrt{(\lambda x_1)^2 + (\lambda x_2)^2}, \sqrt{(\lambda x_2)^2 + (\lambda x_3)^2}, \sqrt{(\lambda x_2)^2 + (\lambda x_1)^2}\right)^T = |\lambda|\, \boldsymbol{A}(\boldsymbol{x}) 。$$

显然，当 $\boldsymbol{x} \neq 0$ 且 $\lambda < 0$ 时，$\boldsymbol{A}(\lambda \boldsymbol{x}) \neq \lambda \boldsymbol{A}(\boldsymbol{x})$，因此 $\boldsymbol{A}$ 不是线性变换。

定义 5.2.2 设 $\boldsymbol{A}$ 是线性空间 $\mathbf{U}$ 到 $\mathbf{V}$ 的线性变换，$\boldsymbol{B}$ 是线性空间 $\mathbf{V}$ 到 $\mathbf{W}$ 上的线性变换，称复合变换

$$\boldsymbol{B}(\boldsymbol{A}(\boldsymbol{x})), \quad \boldsymbol{x} \in \mathbf{U},$$

为 $\boldsymbol{B}$ 和 $\boldsymbol{A}$ 的**乘积变换**，记为 $\boldsymbol{BA}$。

显然 $\boldsymbol{BA}$ 是 $\mathbf{U}$ 到 $\mathbf{W}$ 上线性变换。

定义 5.2.3 设 $\boldsymbol{A}$ 是 $\mathbf{U}$ 上的线性变换，若存在 $\mathbf{U}$ 上的线性变换 $\boldsymbol{B}$ 使得

$$\boldsymbol{BA}(\boldsymbol{x}) = \boldsymbol{x}，\boldsymbol{AB}(\boldsymbol{x}) = \boldsymbol{x}, \quad \boldsymbol{x} \in \mathbf{U},$$

即 $\boldsymbol{BA}$ 和 $\boldsymbol{AB}$ 都是恒等变换，则称 $\boldsymbol{A}$ 是**可逆变换**，$\boldsymbol{B}$ 称为 $\boldsymbol{A}$ 的**逆变换**，记为

$$\boldsymbol{B} = \boldsymbol{A}^{-1}。$$

线性变换有下列性质（其证明作为习题留给读者）：

定理 5.2.1 设 $\boldsymbol{A}$ 是线性空间 $\mathbf{U}$ 到 $\mathbf{V}$ 上的任意一个线性变换，则成立

（1）$\boldsymbol{A}(\mathbf{0}) = \mathbf{0}$，$\boldsymbol{A}(-\boldsymbol{x}) = -\boldsymbol{A}(\boldsymbol{x})$;

（2）若 $\{\boldsymbol{a}_j\}_{j=1}^k$ 是 $\mathbf{U}$ 中一组线性相关的向量，则 $\{\boldsymbol{A}\boldsymbol{a}_j\}_{j=1}^k$ 也是 $\mathbf{V}$ 中一组线性相关的向量;

（3）将 $\mathbf{U}$ 中所有向量在线性变换 $\boldsymbol{A}$ 下的象记为 $\boldsymbol{A}(\mathbf{U})$，即

$$\boldsymbol{A}(\mathbf{U}) = \{\boldsymbol{y} \in \mathbf{V} \mid \boldsymbol{y} = \boldsymbol{A}(\boldsymbol{x}), \boldsymbol{x} \in \mathbf{U}\},$$

则 $\boldsymbol{A}(\mathbf{U})$ 是 $\mathbf{V}$ 的线性子空间（称为 $\boldsymbol{A}$ 的象空间）;

（4）将 $\mathbf{V}$ 中零向量在线性变换 $\boldsymbol{A}$ 下的原象记为 $\mathbf{N}(\boldsymbol{A})$，即

$$\mathbf{N}(\boldsymbol{A}) = \{\boldsymbol{x} \in \mathbf{U} \mid \boldsymbol{A}(\boldsymbol{x}) = \mathbf{0}\},$$

则 $\mathbf{N}(\boldsymbol{A})$ 是 $\mathbf{U}$ 的线性子空间（称为 $\boldsymbol{A}$ 的核空间）。

注意，在定理 5.2.1 的（2）中，若 $\{\boldsymbol{a}_j\}_{j=1}^k$ 是 $\mathbf{U}$ 中一组线性无关的向量，则 $\{\boldsymbol{A}(\boldsymbol{a}_j)\}_{j=1}^k$ 不一定是 $\mathbf{V}$ 中一组线性无关的向量，事实上零变换就是这样。

例 5.2.5 线性变换 $\boldsymbol{A}$： $\mathbf{R}^3 \to \mathbf{R}^2$ 定义为

$$\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{A}(x_1, x_2, x_3) = (x_1 + 2x_2, 2x_2 + x_3)^T,$$

求 $\mathbf{N}(\boldsymbol{A})$ 和 $\boldsymbol{A}(\mathbf{R}^3)$。

解 $\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{0}$ 等价于

$$\begin{cases} x_1 + 2x_2 = 0, \\ 2x_2 + x_3 = 0. \end{cases}$$

其解为

$$\boldsymbol{x} = c(1, -1/2, 1)^T,$$

其中c为任意常数。因此

$$\mathbf{N}(\boldsymbol{A})=\{\,c(1,-1/2,1)^T \mid c\in\mathbf{R}\}。$$

对于任意$(y_1,y_2)^T\in\mathbf{R}^2$，由于线性方程组

$$\begin{cases}x_1+2x_2=y_1,\\ 2x_2+x_3=y_2\end{cases}$$

的增广矩阵$\begin{pmatrix}1&2&0&y_1\\0&2&1&y_2\end{pmatrix}$与系数矩阵$\begin{pmatrix}1&2&0\\0&2&1\end{pmatrix}$的秩皆为 2，所以它有解。这说明$\boldsymbol{A}$为满射，即$\boldsymbol{A}(\mathbf{R}^3)=\mathbf{R}^2$。

下面讨论线性变换与矩阵的关系（下面的所提到的$\mathbf{R}^n$中的向量，皆指列向量）。由例 5.2.1 不难推断，任意一个$n\times m$矩阵A，必确定$\mathbf{R}^m$到$\mathbf{R}^n$上的一个线性变换$\boldsymbol{A}$。事实上，这个线性变换$\boldsymbol{A}$可以如下定义：

$$\boldsymbol{A}(\boldsymbol{x})=A\boldsymbol{x},\quad \boldsymbol{x}\in\mathbf{R}^m。$$

反之，若$\boldsymbol{A}$是$\mathbf{R}^m$到$\mathbf{R}^n$上的线性变换，则存在$n\times m$矩阵A，使得

$$\boldsymbol{A}(\boldsymbol{x})=A\boldsymbol{x},\quad \boldsymbol{x}\in\mathbf{R}^m。$$

事实上，若分别记$\mathbf{R}^m$和$\mathbf{R}^n$上的自然基为$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m\}$和$\{\tilde{\boldsymbol{e}}_1,\tilde{\boldsymbol{e}}_2,\cdots,\tilde{\boldsymbol{e}}_n\}$。因为$\boldsymbol{A}(\boldsymbol{e}_i)\in\mathbf{R}^n$，记

$$\boldsymbol{A}(\boldsymbol{e}_i)=a_{1i}\tilde{\boldsymbol{e}}_1+a_{2i}\tilde{\boldsymbol{e}}_2+\cdots+a_{ni}\tilde{\boldsymbol{e}}_n=(\tilde{\boldsymbol{e}}_1,\tilde{\boldsymbol{e}}_2,\cdots,\tilde{\boldsymbol{e}}_n)\begin{pmatrix}a_{1i}\\a_{2i}\\\vdots\\a_{ni}\end{pmatrix}=\begin{pmatrix}a_{1i}\\a_{2i}\\\vdots\\a_{ni}\end{pmatrix}\quad（i=1,2,\cdots,m），$$

并记$n\times m$矩阵

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1m}\\a_{21}&a_{22}&\cdots&a_{2m}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nm}\end{pmatrix}。$$

则对于$\boldsymbol{x}=(x_1,x_2,\cdots,x_m)^T=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots+x_m\boldsymbol{e}_m\in\mathbf{R}^m$，有

$$\begin{aligned}\boldsymbol{A}(\boldsymbol{x})&=\boldsymbol{A}(x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots+x_m\boldsymbol{e}_m)\\&=x_1\boldsymbol{A}(\boldsymbol{e}_1)+x_2\boldsymbol{A}(\boldsymbol{e}_2)+\cdots+x_m\boldsymbol{A}(\boldsymbol{e}_m)\\&=(\boldsymbol{A}(\boldsymbol{e}_1),\boldsymbol{A}(\boldsymbol{e}_2),\cdots,\boldsymbol{A}(\boldsymbol{e}_m))\boldsymbol{x}=A\boldsymbol{x}.\end{aligned}$$

那么，一般有限维线性空间之间的线性变换与矩阵有什么联系呢？

设$\{\boldsymbol{a}_i\}_{i=1}^m$和$\{\boldsymbol{b}_j\}_{j=1}^n$分别是$m$维线性空间$\mathbf{U}$和$n$维线性空间$\mathbf{V}$中的一组基，$\boldsymbol{A}$是$\mathbf{U}$到$\mathbf{V}$上的线性变换。设$\mathbf{U}$中向量$\boldsymbol{x}$用$\{\boldsymbol{a}_i\}_{i=1}^m$表示的形式为

$$\boldsymbol{x}=\alpha_1\boldsymbol{a}_1+\alpha_2\boldsymbol{a}_2+\cdots+\alpha_m\boldsymbol{a}_m,$$

两边作用线性变换$\boldsymbol{A}$，由线性变换的性质得，

$$\boldsymbol{A}(\boldsymbol{x})=\alpha_1\boldsymbol{A}(\boldsymbol{a}_1)+\alpha_2\boldsymbol{A}(\boldsymbol{a}_2)+\cdots+\alpha_m\boldsymbol{A}(\boldsymbol{a}_m)。$$

这就是说，线性变换由其对一组基的变换规律完全决定。

由于对于$i=1,2,\cdots,m$，$\boldsymbol{A}(\boldsymbol{a}_i)\in\mathbf{V}$，因此它可以用基$\{\boldsymbol{b}_j\}_{j=1}^n$线性表示，记

$$A(\boldsymbol{a}_i)=(\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_n)\begin{pmatrix}a_{1i}\\a_{2i}\\\vdots\\a_{ni}\end{pmatrix},\quad i=1,2,\cdots,m。$$

于是，若记

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1m}\\a_{21}&a_{22}&\cdots&a_{2m}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nm}\end{pmatrix},$$

则

$$\begin{aligned}&(\boldsymbol{A}(\boldsymbol{a}_1),\boldsymbol{A}(\boldsymbol{a}_2),\cdots,\boldsymbol{A}(\boldsymbol{a}_m))\\&=(\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_n)\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1m}\\a_{21}&a_{22}&\cdots&a_{2m}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nm}\end{pmatrix}=(\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_n)A。\end{aligned}$$

所以

$$\begin{aligned}\boldsymbol{A}(\boldsymbol{x})&=\alpha_1\boldsymbol{A}(\boldsymbol{a}_1)+\alpha_2\boldsymbol{A}(\boldsymbol{a}_2)+\cdots+\alpha_m\boldsymbol{A}(\boldsymbol{a}_m)\\&=(\boldsymbol{A}(\boldsymbol{a}_1),\boldsymbol{A}(\boldsymbol{a}_2),\cdots,\boldsymbol{A}(\boldsymbol{a}_m))\begin{pmatrix}\alpha_1\\\alpha_2\\\vdots\\\alpha_m\end{pmatrix}=(\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_n)\ A\begin{pmatrix}\alpha_1\\\alpha_2\\\vdots\\\alpha_m\end{pmatrix}。\end{aligned}$$

这就是说，线性变换 $\boldsymbol{A}$ 由 $n\times m$ 矩阵 $A=\left(a_{ij}\right)_{n\times m}$ 唯一确定，称 A 为线性变换 $\boldsymbol{A}$ 在基 $\{\boldsymbol{a}_i\}_{i=1}^m$ 和 $\{\boldsymbol{b}_j\}_{j=1}^n$ 下的**表示矩阵**。顺便地，我们得到：当 $\boldsymbol{x}$ 在基 $\{\boldsymbol{a}_i\}_{i=1}^m$ 下的坐标为 $(\alpha_1,\alpha_2,\cdots,\alpha_m)^T$ 时，$\boldsymbol{A}(\boldsymbol{x})$ 在基 $\{\boldsymbol{b}_j\}_{j=1}^n$ 下的坐标便是 $A(\alpha_1,\alpha_2,\cdots,\alpha_m)^T$。

特别地，当 $\mathbf{U}$ 和 $\mathbf{V}$ 分别为 $\mathbf{R}^m$ 和 $\mathbf{R}^n$ 时，且基都取为自然基时，便从以上讨论得到：

例 5.2.6 设 $\mathbf{R}^2$ 上的线性变换 $\boldsymbol{A}$ 将任意给定的向量 $\boldsymbol{x}$ 绕原点逆时针旋转角度 θ，求 $\boldsymbol{A}$ 在自然基 $\{\boldsymbol{e}_1，\boldsymbol{e}_2\}$ 下的表示矩阵。

解 $\boldsymbol{e}_1$，$\boldsymbol{e}_2$ 绕原点逆时针旋转角度 θ 后，坐标分别是 $\begin{pmatrix}\cos\theta\\\sin\theta\end{pmatrix}$ 和 $\begin{pmatrix}-\sin\theta\\\cos\theta\end{pmatrix}$（见图 5.2.6），即

$$(\boldsymbol{A}(\boldsymbol{e}_1),\boldsymbol{A}(\boldsymbol{e}_2))=(\boldsymbol{e}_1,\boldsymbol{e}_2)\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}。$$

所以，旋转变换 $\boldsymbol{A}$ 在自然基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2\}$ 下的表示矩阵为

$$A=\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}。$$

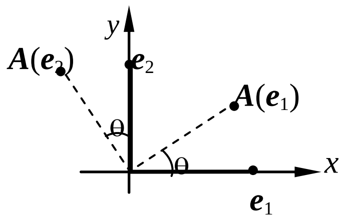

图 5.2.6

由于任意向量 $\boldsymbol{x}=\begin{pmatrix}x\\y\end{pmatrix}$ 在自然基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2\}$ 下的坐标就是它的分量，因此它经过变换 $\boldsymbol{A}$ 后，$\boldsymbol{A}(\boldsymbol{x})$ 在 $\{\boldsymbol{e}_1,\boldsymbol{e}_2\}$

下的坐标是

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

这就是例 5.2.1（4）的结果。

三．不同基下表示矩阵的关系

为叙述简洁起见，下面只讨论线性空间到其自身的线性变换。

设 $\boldsymbol{A}$ 是 m 维线性空间 $\mathbf{U}$ 上的线性变换，$\{\boldsymbol{a}_i\}_{i=1}^m$ 是 $\mathbf{U}$ 的一组基。由前面所述，$\boldsymbol{A}$ 在这组基下的表示矩阵是指满足

$$(\boldsymbol{A}(\boldsymbol{a}_1), \boldsymbol{A}(\boldsymbol{a}_2), \cdots, \boldsymbol{A}(\boldsymbol{a}_m)) = (\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m)A$$

的矩阵 A，它是一个 m 阶方阵。作为前面讨论的特例可知：如果在基 $\{\boldsymbol{a}_i\}_{i=1}^m$ 下，$\boldsymbol{x}$ 的坐标为 $(\alpha_1, \alpha_2, \cdots, \alpha_m)^T$，则 $\boldsymbol{A}(\boldsymbol{x})$ 的坐标是 $A(\alpha_1, \alpha_2, \cdots, \alpha_m)^T$。

显然，有限维线性空间上的恒等变换在任意基下的表示矩阵都是单位矩阵；零变换在任意基下的表示矩阵都是零矩阵。

下面的定理说明了有限维线性空间上的线性变换在不同的基下的表示矩阵的关系。

定理 5.2.2　设 $\boldsymbol{A}$ 是 m 维线性空间 $\mathbf{U}$ 上的任意一个线性变换，$\{\boldsymbol{a}_i\}_{i=1}^m$ 和 $\{\boldsymbol{b}_i\}_{i=1}^m$ 是 $\mathbf{U}$ 的两组基，从 $\{\boldsymbol{a}_i\}_{i=1}^m$ 到 $\{\boldsymbol{b}_i\}_{i=1}^m$ 的过渡矩阵为 T。若 $\boldsymbol{A}$ 在基 $\{\boldsymbol{a}_i\}_{i=1}^m$ 和 $\{\boldsymbol{b}_i\}_{i=1}^m$ 下的表示矩阵分别是 A 和 B，则成立

$$B = T^{-1}AT。$$

证　$\boldsymbol{A}$ 在基 $\{\boldsymbol{a}_i\}_{i=1}^m$ 和 $\{\boldsymbol{b}_i\}_{i=1}^m$ 下的表示矩阵分别是 A 和 B，即

$$(\boldsymbol{A}(\boldsymbol{a}_1), \boldsymbol{A}(\boldsymbol{a}_2), \cdots, \boldsymbol{A}(\boldsymbol{a}_m)) = (\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m)A,$$

和

$$(\boldsymbol{A}(\boldsymbol{b}_1), \boldsymbol{A}(\boldsymbol{b}_2), \cdots, \boldsymbol{A}(\boldsymbol{b}_m)) = (\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_m)B,$$

而从 $\{\boldsymbol{a}_i\}_{i=1}^m$ 到 $\{\boldsymbol{b}_i\}_{i=1}^m$ 的过渡矩阵为 T，即

$$(\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_m) = (\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m)T。$$

对每个 $\boldsymbol{b}_i$ 作线性变换 $\boldsymbol{A}$，利用 $\boldsymbol{A}$ 的线性性质，并注意到 T 是可逆矩阵，由上式得

$$\begin{aligned}(\boldsymbol{A}(\boldsymbol{b}_1), \boldsymbol{A}(\boldsymbol{b}_2), \cdots, \boldsymbol{A}(\boldsymbol{b}_m)) &= (\boldsymbol{A}(\boldsymbol{a}_1), \boldsymbol{A}(\boldsymbol{a}_2), \cdots, \boldsymbol{A}(\boldsymbol{a}_m))T \\ &= (\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m)AT = (\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_m)T^{-1}AT,\end{aligned}$$

因此

$$B = T^{-1}AT。$$

证毕

下表列出在 $\mathbf{U}$ 上的线性变换 $\boldsymbol{A}$ 下，$\mathbf{U}$ 的向量 $\boldsymbol{x}$ 的坐标变化情况：

	在基 $\{\boldsymbol{a}_i\}_{i=1}^m$ 下的坐标	在基 $\{\boldsymbol{b}_i\}_{i=1}^m$ 下的坐标
$\boldsymbol{x}$	$\boldsymbol{\xi}$	$T^{-1}\boldsymbol{\xi}$
$\boldsymbol{A}(\boldsymbol{x})$	$A\boldsymbol{\xi}$	$T^{-1}A\boldsymbol{\xi}$

例 5.2.7　在 $\mathbf{P}_3$ 中考虑求导运算 $\boldsymbol{A} = \dfrac{d}{dx}$。由于在基 $\{1, x, x^2, x^3\}$ 下，

$$(\boldsymbol{A}(1),\ \boldsymbol{A}(x),\ \boldsymbol{A}(x^2),\ \boldsymbol{A}(x^3)) = (0, 1, 2x, 3x^2)$$

$$=(1,x,x^2,x^3)\begin{pmatrix}0&1&&\\&0&2&\\&&0&3\\&&&0\end{pmatrix}=(1,x,x^2,x^3)A,$$

以及

$$\left(1,x,\frac{3x^2-1}{2},\frac{5x^3-x}{2}\right)=(1,x,x^2,x^3)\begin{pmatrix}1&0&-\frac{1}{2}&0\\0&1&0&-\frac{1}{2}\\0&0&\frac{3}{2}&0\\0&0&0&\frac{5}{2}\end{pmatrix}=(1,x,x^2,x^3)T,$$

因此，它在基$\left\{1,x,\frac{3x^2-1}{2},\frac{5x^3-x}{2}\right\}$下的表示矩阵应为

$$B=T^{-1}AT=\begin{pmatrix}0&1&0&2\\0&0&3&0\\0&0&0&5\\0&0&0&0\end{pmatrix}。$$

可以直接计算

$$(\boldsymbol{A}(1),\ \boldsymbol{A}(x),\ \boldsymbol{A}\left(\frac{3x^2-1}{2}\right),\ \boldsymbol{A}\left(\frac{5x^3-x}{2}\right))=\left(0,1,3x,\frac{1}{2}(15x^2-1)\right)$$

$$=\left(1,x,\frac{3x^2-1}{2},\frac{5x^3-x}{2}\right)\begin{pmatrix}0&1&0&2\\0&0&3&0\\0&0&0&5\\0&0&0&0\end{pmatrix}。$$

我们已经知道，$\mathbf{P}_3$ 中元素 $p(x)=4x^3+3x^2+2x+1$在基$\{1,x,x^2,x^3\}$下的坐标是 $\boldsymbol{\xi}=(1,2,3,4)^T$，则在$\left\{1,x,\frac{3x^2-1}{2},\frac{5x^3-x}{2}\right\}$下的坐标是$T^{-1}\boldsymbol{\xi}=\frac{2}{5}(5,7,5,4)^T$。在求导后，$\boldsymbol{A}(p)$ 在$\{1,x,x^2,x^3\}$下的坐标为

$$A\boldsymbol{\xi}=\begin{pmatrix}0&1&&\\&0&2&\\&&0&3\\&&&0\end{pmatrix}\begin{pmatrix}1\\2\\3\\4\end{pmatrix}=\begin{pmatrix}2\\6\\12\\0\end{pmatrix},$$

则在$\left\{1,x,\frac{3x^2-1}{2},\frac{5x^3-x}{2}\right\}$下的坐标为

$$T^{-1}A\boldsymbol{\xi}=B(T^{-1}\boldsymbol{\xi})=\frac{2}{5}\begin{pmatrix}0&1&0&2\\0&0&3&0\\0&0&0&5\\0&0&0&0\end{pmatrix}\begin{pmatrix}5\\7\\5\\4\end{pmatrix}=\begin{pmatrix}6\\6\\8\\0\end{pmatrix}。$$

另一方面，通过直接计算有

$$\boldsymbol{A}(p)=(4x^3+3x^2+2x+1)'=12x^2+6x+2=8\frac{3x^2-1}{2}+6x+6,$$

这与利用坐标向量计算的结果一致。

四．相似矩阵与相似变换

定义 5.2.4 设 A 和 B 是同阶方阵，若存在同阶可逆方阵 T 使得

$$B=T^{-1}AT,$$

则称 A 和 B 是**相似矩阵**（简称 A 和 B **相似**），记作 $A\sim B$。

显然，相似矩阵具有相同的行列式。将 A 变为 $T^{-1}AT$ 称为对 A 作**相似变换**。

定理 5.2.2 告诉我们，线性空间 $\mathbf{U}$ 上同一个线性变换在任意两组基下的表示矩阵必定相似。反过来，可以证明，如果线性空间上线性变换 $\boldsymbol{A}$ 在一组基下的表示矩阵为 A，矩阵 B 与 A 相似，则必有线性空间的另一组基，$\boldsymbol{A}$ 在这组基下的表示矩阵为 B。

那么很自然地要问，对一个给定的线性变换 $\boldsymbol{A}$，能否找出 $\mathbf{U}$ 的一组基，使 $\boldsymbol{A}$ 在这组基下的表示矩阵尽可能简单？由于最简单的表示矩阵是对角阵，因此上面的问题等价于：对于任意一个给定的方阵，能否找到同阶可逆方阵 T，使得 $T^{-1}AT$ 是对角阵，或者至少是块对角阵？这是我们在下一节要讨论的问题。

五．习　　题

1．（1）、（3），2，3，4，6，7。